Applicants: Graham P. Allaway et al.
Serial No.: 09/904,356
Filed: July 12, 2001
Exhibit 18

:mental Immunology, Fourth Edition,
. . by William E. Paul
Published by Lippincott-Raven Publishers, Philadelphia 1999.

CHAPTER 5

Immunoglobulins: Molecular Genetics

Edward E. Max

The study of the molecular biology of immunoglobulin (Ig) genes represents one of the first triumphs of recombinant DNA technology. Before the advent of gene cloning, Ig genes could be studied only indirectly by inferences from amino acid sequences. Many perplexing questions were resolved when it became possible to examine directly the genes themselves. Recently the cloned genes have moved beyond the pure research laboratory to be used as tools for various applied engineering projects. This chapter summarizes some of these exciting advances in both the basic and applied arenas.

The unique mystery of antibody genes lies in the diversity of proteins they encode. This diversity exists at several levels.

Most striking is the diversity of antigen-combining sites of these molecules. The classic studies of Landsteiner suggested that the repertoire of binding specificities of antibodies is essentially unlimited. The diversity of binding specificities is explained by the diversity of amino acid sequences found in the N-terminal domain of both light (L) and heavy (H) chains—the variable (V) region—each containing three regions of especially high variability (hypervariable regions) which correspond to the loops of the protein that contact antigen, or complementarity determining regions (CDRs), as discussed in Chapter 3. Yet on the C-terminal end, the single domain of the L chain and the three (or four depending on isotype) domains of H chains were found to be invariant within each class of L or H chains; these segments are designated constant (C) regions. Many models were proposed to explain the unprecedented diversity found in Ig V regions. One extreme model suggested that the immense diversity of V regions was directly encoded in the germline genome, presumably a result of gene duplication and mutation acting over evolutionary time. At the other extreme, the somatic mutation model supposed that very few V-region sequences were encoded in the genome and that a special somatic mutation mechanism operated on these sequences to increase diversity within the life span of the organism. Regardless of whether sequence diversification occurred in phylogeny (germline diversity) or ontogeny (somatic mutation), another question remained: How did the C regions of Ig genes escape such changes? In 1965, Dreyer and Bennett (1) proposed that for each class of Ig genes there might be only a single C-region gene, which was encoded in the germline separately from the multiple V-region genes; in the development of an antibody-producing cell, one of the V-region sequences would become associated with the C-region sequence, leading to a complete (V + C) gene, which the cell could then express. Thus, mechanisms that increase diversity in the isolated V-region genes might leave the single C-region gene at its distant locus untouched. This model, with its proposal of gene rearrangement occurring independently in each lymphocyte, was revolutionary in that it violated the then-accepted notion that DNA is the same in all cells of the organism. Clearly, a definitive assessment of Dreyer and Bennett's proposal

E. E. Max: Laboratory of Cell and Viral Regulation, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, Maryland 20892.

and an evaluation of the relative significance of somatic and germline interpretations of V-region diversity required a direct analysis of the genes in question. Two additional mysteries: given the fact that each B lymphocyte should contain two copies of each gene locus (i.e., from the maternally and paternally derived chromosomes), why does the cell express only a single L chain and H chain, as if the locus on the nonexpressed chromosome were somehow silenced—the phenomenon known as "allelic exclusion"? And how can one explain the fact that affinity of serum antibodies for antigen increases over a period of weeks after antigen exposure—the phenomenon of "affinity maturation"?

Apart from the diversity of V regions in both L and H chains, H chains exhibit a different sort of diversity that also demands a molecular biologic explanation: all developing B cells synthesize IgM initially and can switch H-chain isotype from μ to γ, ε, or α only later in their maturation. As the expressed C-region "switches," the cell continues to express the same L- and H-chain V regions, so that antigen specificity remains unchanged. Thus, in addition to understanding how, in different cells, a single C region can become associated with multiple different V regions (V-C recombination), we need to consider the molecular mechanism by which, during lymphocyte development, a single V region may become associated sequentially with several C regions (H-chain switch).

A final level of diversity exhibited by Ig H chains is represented by the alternative forms of Ig found embedded in the membrane of B cells versus those in blood and secretions. Membrane Igs have C-terminal extensions containing hydrophobic amino acids that associate with membrane lipids, whereas secreted Igs lack this C-terminal piece but are otherwise identical to the membrane counterparts. Analysis of Ig genes has shown how these two forms are encoded in the genome.

This chapter will begin with a brief discussion of V gene assembly in H- and L-chain genes. We then describe the H-chain locus—including molecular explanations for the membrane forms of Ig and isotype switching—followed by descriptions of κ and λ gene loci; however, a detailed discussion of each germline V-gene locus is deferred until later in the chapter. Next we consider in detail the DNA recombination events underlying V-gene assembly and the regulation of this process to maintain allelic exclusion. The chapter continues with a discussion of the mechanisms contributing to V-region diversity: the germline V repertoire, junctional diversity, and somatic mutation. A discussion of the regulation of Ig gene expression follows. The chapter ends with several topics in the "applied science" of Ig genes.

The investigations described in this chapter have been chosen from the literature to facilitate a clear exposition of the important issues rather than to provide a comprehensive compendium of data and references on Ig genes. In these descriptions, most of the discussion focuses on murine and human Ig genes. Murine genes were studied first because of the availability of pristane-induced murine myelomas of BALB/c mice, which served as convenient monoclonal sources of Ig protein for early structural studies. The same myelomas then provided messenger RNA (mRNA) and DNA for molecular biology analysis, which was greatly facilitated by the fact that these myelomas derived from the same genetic background—the inbred BALB/c strain. Later, study of the homologous human loci showed many fundamental similarities between the Ig genes of these two species, whereas some other mammalian orders show surprisingly significant differences.

Isotype switching and somatic mutation of Ig genes are covered in more detail in separate chapters of this text (Chapters 23 and 24).

OVERVIEW OF IMMUNOGLOBULIN V-GENE ASSEMBLY

In the late 1970s, experiments on L-chain genes established that the Dreyer-Bennett hypothesis was fundamentally correct: each lymphocyte expresses only a single Ig molecule encoded by one VL and one VH gene, each having been "activated" by a recombination event that brings the V gene near its respective C-region gene. This conclusion was supported by comparisons of Ig genes from B-lymphoid cells, particularly murine myelomas, and the corresponding gene loci from "germline" DNA. (Although true germline DNA can experimentally be obtained only from sperm, any nonlymphoid DNA is assumed to be representative of germline DNA because the rearrangements of Ig genes occur only in lymphoid cells. When DNA from sperm versus other nonlymphoid tissues has been compared by Southern blots, the results have been identical. Therefore, despite the risk of some imprecision, nonlymphoid DNA samples are conventionally referred to as germline whether the DNA is from sperm, whole embryo, liver, placenta, or other nonlymphoid sources.)

Evidence from Southern Blots and Gene Cloning

Initially the myeloma and germline DNA samples were compared by Southern blotting using hybridization probes derived from myeloma complementary DNA (cDNA). As schematically shown in Fig. 1 for an analysis of κ L-chain genes, a Cκ probe detects only a single band in germline DNA, consistent with a single Cκ gene. A probe representing an expressed Vκ gene detects several bands, as though hybridizing to a family of related sequences. Moreover, although not shown in Fig. 1, probes representing different expressed Vκ genes are found to hybridize to a different set of bands, representing a different family of related Vκ genes. These observations support the hypothesis of multiple V genes, single C gene. The novel recombination postulate of the Dreyer-Bennett hypothesis is supported by the differences observed when these probes are hybridized to myeloma DNA instead of germline DNA. As shown in Fig. 1, the recombination bringing a V gene close to a C gene can cause an alteration in size of the Cκ-hybridizing restriction fragment. The new rearranged band may be larger, smaller, or fortuitously the same size as the germline band, depending on the location of the restriction sites flanking the V and C genes. One of the V-region bands may similarly be expected to be rearranged in the myeloma so as to lie on a different-sized fragment, the same fragment that hybridizes to the Cκ probe. Results like these for κ and λ genes strongly supported the Dreyer-Bennett hypothesis and forcefully challenged the concept that every cell in the body has identical genes (2,3). In panels of myelomas analyzed for Cκ recombination by Southern blotting, many showed evidence of DNA rearrangement on both allelic chromosomes. This result argued against the possibility that allelic exclusion might be explained by a mechanism that allowed recombination on only one chromosome, and it raised questions about the nature of the "second" gene rearrangement in these cells, as discussed later in this chapter.

A more complete understanding of recombination of Ig genes developed from sequence analysis of cloned myeloma versus germline DNA. The general structures of the germline V genes are similar for the three Ig loci: H chain, κ, and λ. Each V gene begins with a sequence encoding a signal peptide of about 22 amino acids. (Signal peptides are found at the N-terminus of most proteins destined for secretion or expression on the cell membranes; after rout-

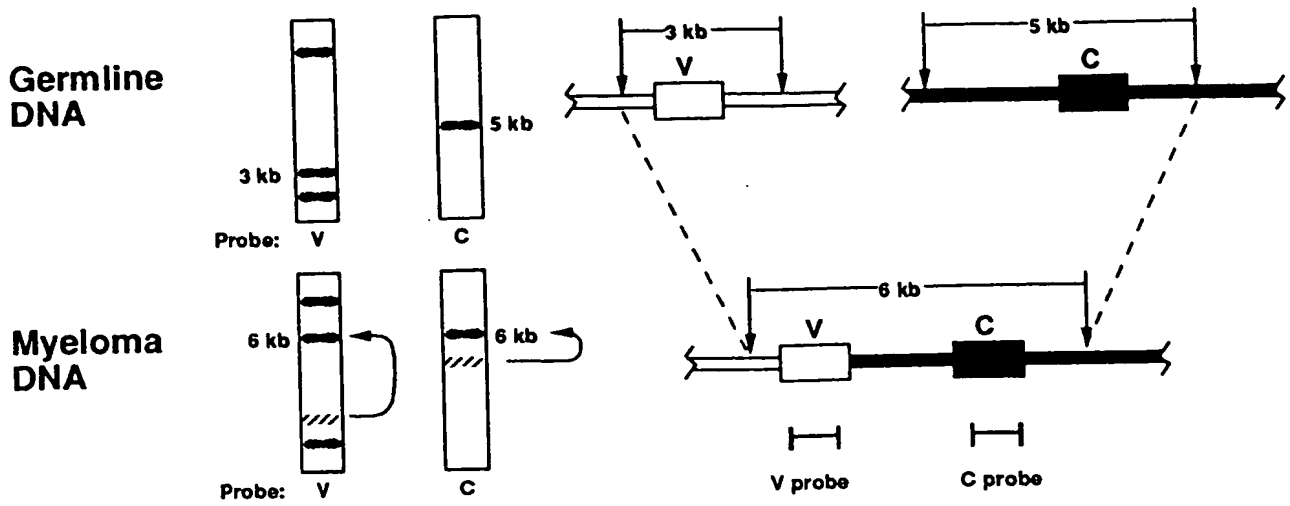


FIG. 1. Southern blot demonstration of rearrangement of Ig V and C region genes. *Eco*RI sites in this hypothetical example are indicated by arrows. In germline DNA (*upper drawings*) V and C are an unknown distance apart and are found by Southern blotting (*left*) to lie on *Eco*RI fragments of 3 and 5 kb, respectively. The V region probe hybridizes to a family of related genes (shown by bands above and below the 3-kb band). In myeloma DNA (*lower panel*), V and C genes have been brought into close proximity and, in this example, are no longer separated by *Eco*RI sites; both genes are found on the same *Eco*RI fragment of 6 kb, which is thus identified by either probe. The germline-sized fragments (hatched bands in the Southern blots) may or may not be preserved in the myeloma, depending on whether the nonexpressed homologous chromosome has remained in its germline (unrearranged) state. In many myelomas both chromosomes are present and both are rearranged.

ing the protein to the endoplasmic reticulum, the peptide is generally removed by specific peptidases.) Within codon 4 (numbering backward from the beginning of the mature protein sequence), the coding sequence is interrupted by an intron, usually 0.1 to 0.3 kb long. What was unanticipated was the discovery that each V-region gene as it exists in the germline is incomplete, and that recombination is necessary to assemble a complete V gene (4). For example, most murine κ chains have V regions 108 amino acids in length, but murine germline Vκ genes encode only about 95 of these. The remaining 13 amino acids are encoded by segments known as J (joining) regions that lie upstream of the C-region gene (5,6). An assembled Vκ gene thus results from recombination that joins one of many germline Vκ genes to one of five Jκ gene segments (Fig. 2A). A similar recombination event is necessary to assemble a complete Vλ-chain sequence from germline Vλ and Jλ genes (7). For H chains, recombination assembles a V region from three types of germline elements; between the residues encoded by germline VH and JH elements there are interposed variable numbers of amino acids—commonly from zero to eight residues—encoded by a D (diversity) region. The assembly of a complete H-chain V region occurs in two separate steps (Fig. 2B): initially one of several germline DH regions joins with one of the JH regions; then a germline VH region is added to complete the assembled VDJ H-chain gene.

How Recombination Contributes to Diversity

The V-assembly recombination contributes in two significant ways to the diversity of antigen-binding specificities. First, because there are multiple germline V regions and multiple D and J regions, the number of possible combinations of Vλ, Jλ, VH, DH, and JH is the multiplication product of the numbers of each of these five classes of germline sequence elements. This repertoire is vastly larger than could be achieved by devoting the same total lengths of DNA sequence to preassembled V regions. A second factor that increases diversity was recognized by comparing nucleotide sequences of various myeloma genes to their germline precursors. For example, as shown in Fig. 3A, a comparison between the Vκ gene expressed in the murine myeloma MOPC41 and the corresponding germline Vκ and Jκ genes shows that the myeloma gene matches the germline precursor through the second nucleotide of codon 95; the VJ recombination junction clearly occurs at this point because sequence beyond this position in the myeloma gene clearly derives from Jκ1. Similar analyses of other myelomas show that the recombination junctions can occur at several different positions within codon 95 or 96. As shown in Fig. 3B, this flexibility of the position of the recombination junction increases the diversity of the affected codons. H-chain V regions exhibit this flexibility at both VD and DJ junctions. In addition, many H-chain VDJ junctions (and a smaller percentage of L-chain VJ junctions) show insertions of a few extra nucleotides not present in the germline precursors; the mechanism of these insertions will be discussed later in this chapter. Significantly, the three-dimensional structure of Igs established from x-ray crystallography shows that the Vλ-JL junction and the VH-DH-JH junction both form CDR3 loops that can contact antigen; thus this junctional diversity is physiologically relevant for diversifying antigen binding. The important role of D junctional amino acids for antigen binding has been verified by mutational analysis (8). In addition, many H-chain VDJ junctions (and a smaller percentage of L-chain VJ junctions) reveal insertions of a few extra nucleotides not present in the germline precursors; the mechanism of these insertions—known as N regions—will be discussed later in this chapter.

When the flexibility of the position of recombination was initially discovered, it was hard to understand how the germline elements could be joined with such variability and yet maintain the correct triplet reading frame between V and J. (An out-of-frame recombi-

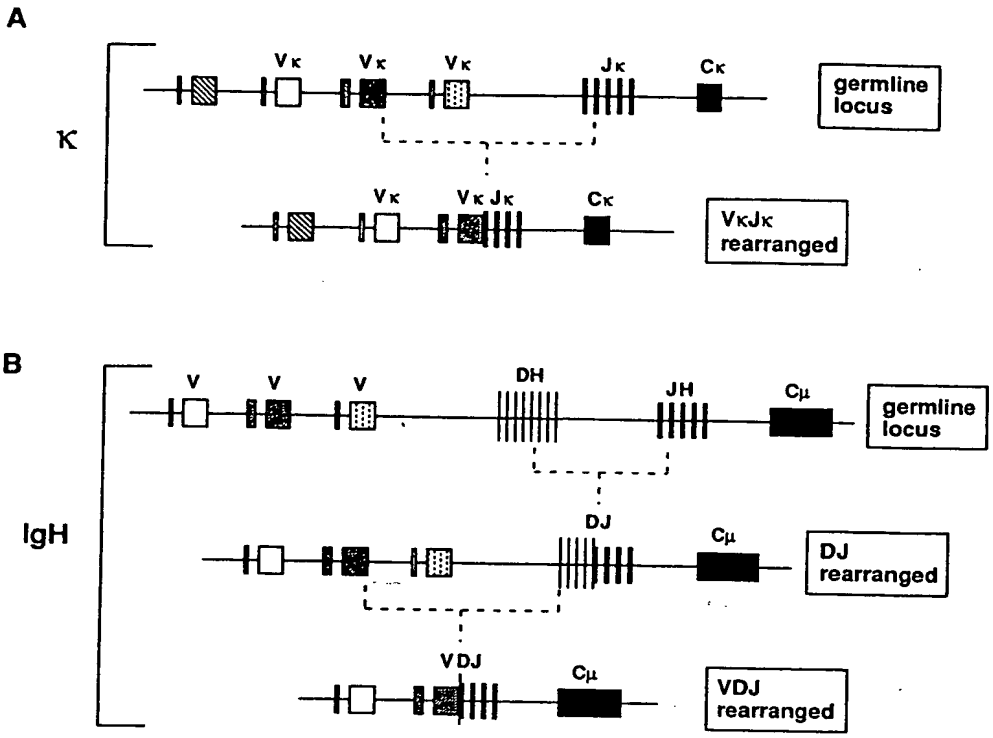


FIG. 2. V assembly recombination. **A:** In the κ locus a single recombination event joins a germline Vκ region with one of the Jκ segments. **B:** In the IgH locus an initial recombination joins a D segment to a JH segment. A second recombination completes the V assembly by joining a VH to DJH.

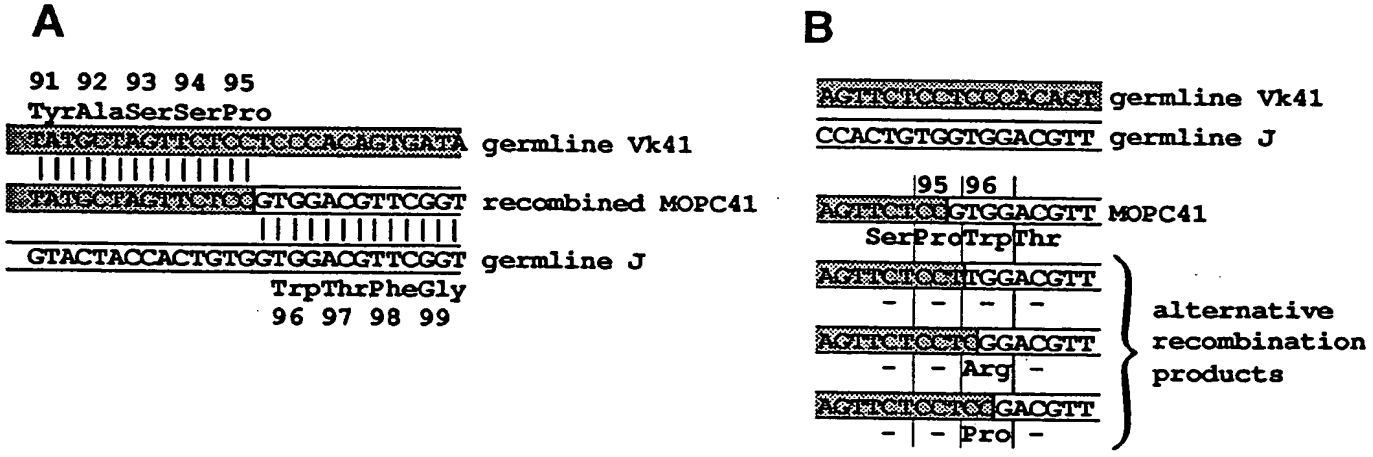


FIG. 3. Vκ-Jκ recombination at single base resolution. **A:** The sequence of the recombined MOPC41 κ gene around the VJ junction is shown at center, with the sequences of the two germline precursors (Vκ41 and Jκ1) shown above and below. The germline origins of the recombined gene are indicated by the vertical lines and the shading of the V-derived sequence. **B:** The consequences of joining the same germline sequences (from **A**) at four different positions are shown. Of the four alternative recombination products illustrated, the top one is that actually found in MOPC41. The second example has a single nucleotide difference but no change in encoded amino acid sequence. The third and fourth alternatives yield Arg or Pro at position 96; both of these amino acids have been found at this position in sequenced mouse κ chains.

nation would cause the entire C region to be read in a nonsense reading frame, so the gene would be nonfunctional.) It soon became clear, however, that if one looks beyond the subset of assembled V regions that are expressed in myeloma antibodies—a subset selected for expression of a functional L and H chain—one can find many assembled V genes with out-of-frame recombination junctions (9). Indeed, in unselected VJ recombinations the frequency of in-frame junctions is about 1/3, as predicted for a recombination mechanism insensitive to reading frame. In myelomas with rearrangements on both allelic copies of an Ig gene locus, the unexpressed recombination is generally out-of-frame or "non-productive." For H-chain VDJ recombination, one could theoretically retain the correct reading frame between V and J while allowing the interposed D-region segments to be used in all three reading frames. In murine H chains, however, only a single D-region reading frame is generally found, and several mechanisms prevent expression of antibodies with D regions in the other two reading frames (10). In human antibodies this intense selection against variant reading frames is not found (11), allowing for additional sequence diversity. The generation of V-region diversity in the three Ig gene loci (IgH, κ, and λ) is considered in more detail in a later section.

Recombination Signal Elements

Analysis of DNA sequences flanking the germline V-, D-, and J-region sequences showed two conserved sequence elements that apparently play a role in the recombination event signaling the position where the DNA should rearrange. The first signal element is a 7-mer CACTGTG that occurs as a consensus sequence 5′ to the Jκ coding sequences, with its (reverse) complement CACAGTG appearing 3′ to the Vκ coding sequences. The second element is a 9-mer GGTTTTTGT that appears about 23 nucleotides 5′ to the Jκ 7-mer, its complement ACAAAAACC appearing about 12 nucleotides 3′ to the Vκ 7-mer (5,6). The likelihood that these recombination signal sequences (RSS) are significant in the recombination is reinforced by their appearance at similar positions in L- and H-chain Ig genes throughout phylogeny as well as in T-cell receptor (TCR) genes (see Chapter 10), which undergo similar V assembly recombinations; furthermore, there are no other well-conserved sequences flanking these genes. In all of these systems the length of the spacer between the 7-mer and 9-mer appears significant. Recombination apparently occurs only between one coding sequence with a 12-bp spacer and another coding sequence with a 23-bp spacer, a requirement referred to as the 12/23 rule. The benefit of this requirement may be that futile recombinations, such as between two Vκ or two Jκ gene segments, are prevented. Although a computerized alignment of several hundred spacer sequences has detected some preferred nucleotides at specific positions (12), mutations of spacer sequences in plasmid recombination substrates have little effect on recombination frequency. The length of the spacers flanking H- and L-chain V, D, and J elements are shown in Fig. 4.

Although the complementarity of the Vκ and Jκ copies of the 7-mer and 9-mer signal elements initially led to the hypothesis that these elements might participate in the formation of a stem-and-loop intermediate in the recombination reaction, current evidence strongly favors an alternative role for the RSS: as recognition sequences for DNA-binding proteins mediating the recombination. This evidence is presented later in a detailed discussion of V(D)J recombination.

Because of the conservation of the RSS elements among κ, λ, IgH genes, and TCR genes, the enzymatic recombinase machinery that assembles complete V genes from germline precursors is believed to be the same in all these systems. This notion is reinforced by much other evidence, including the observations that germline TCR V-gene segments can be correctly rearranged when introduced into pre-B cells and that hybrid Ig-TCR rearrangement can occur (although only in abnormal cells, as discussed in a later section).

THE THREE IMMUNOGLOBULIN GENE LOCI

This section presents an overview of the three Ig loci: H chain, κ, and λ. The V regions of these loci are described in a later section on germline diversity (except that the tiny murine Vλ repertoire is discussed in the present section).

Heavy-Chain Genes

In the development of a B-lymphocyte, the cell initially produces IgM with a binding specificity determined by the productively rearranged VH and VL regions. Subsequently, each B cell and its

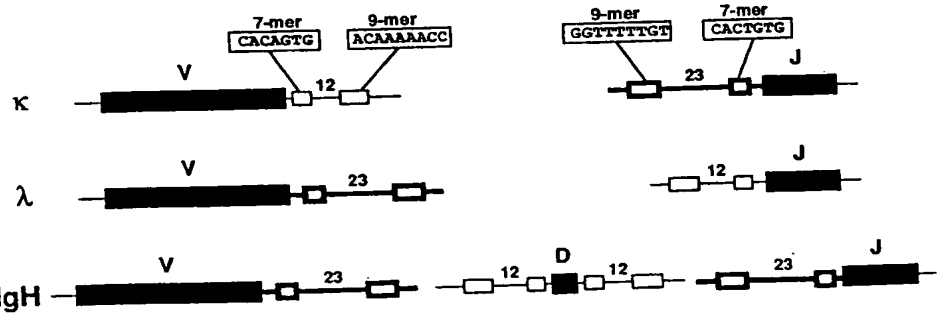


FIG. 4. Conserved elements flank germline V, D, and J region genes. Conserved 7-mer and 9-mer RSS lie adjacent to V, D, and J coding sequences and are important for targeting V(D)J recombination. The 7-mer and 9-mer elements are separated by spacer regions of about 12 bp (illustrated by thin lines in the figure) or 23 bp (thick lines). Depending on the locus, V regions may be flanked by 12- or 23-bp RSS, and similarly for J regions. But one of each type of element must be present for recombination to occur, a requirement that prevents futile recombination events (e.g., J to J).

progeny cells synthesize antibodies with the same L- and H-chain V regions; but they may later switch the isotype of the H chain. Early evidence for this developmental scheme includes (among other observations) (a) the isotype shift seen during the course of an immune response (13); (b) the ability of B-cell clones—myelomas (14,15) and splenic foci (16)—to express IgM plus another isotype, with identical VH regions; and (c) in vivo ablation studies suggesting that IgM-producing cells are the precursors of IgG producers (17). The molecular mechanism by which one part of a protein can change while another part remains unchanged has generated considerable interest.

Several groups (18–20) have demonstrated that active rearranged α, γ2b, and γ1 genes isolated from myelomas expressing the respective H chains contain—between their V and C regions—DNA sequences derived from the DNA upstream of the germline Cμ gene, including one or more JH sequences. These observations led to the model (Fig. 5) that the VH region rearranges initially to a position 5′ to the μ gene (leading to IgM production), and that when a cell expresses a new isotype the Cμ-region gene is replaced by the CH region encoding the new isotype. This isotype switch appears to result from a deletion of the CH genes between the assembled VDJ and the CH gene expressed after the switch. Early support for this deletion model came from analysis of the content of specific CH genes in myelomas that had undergone different switch recombinations. Solution hybridization kinetics or Southern blotting with cDNA-derived CH probes confirmed that switching was associated with loss of CH sequences from the cell. From the specific C regions lost in myelomas expressing different isotypes, it was possible to predict a linear order of the different CH genes on the chromosome (21,22).

A more detailed picture of the H-chain locus emerged as many laboratories reported the isolation of genomic clones for CH genes. In general, these clones were obtained in the early 1980s by screening genomic DNA libraries with cDNA probes derived from myeloma mRNA. From the wealth of data generated, we can consider only a few interesting conclusions because of space limitations.

One striking characteristic of CH genes is that the 100 to 110 amino acid domains—identified by internal homologies of amino acid sequences and by three-dimensional structural analysis (x-ray crystallography)—are encoded as intact exons, separated from other domain segments by introns of 0.1 to 0.3 kb (23–25). Thus, for example, the mouse γ2b protein has three major domains (CH1, CH2, and CH3), with a small hinge domain between CH1 and CH2. The gene structure (23,26,27) may be summarized as follows:

CH1	—intron—	**hinge**	—intron—	**CH2**	—intron—	**CH3**
(292)	(314)	(64)	(106)	(328)	(119)	(322)

where the numbers in parentheses represent the number of nucleotides in each segment. As an interesting contrast, the hinge region of the α gene is encoded contiguously with the CH2 domain with no intervening intron (25), whereas the unusually long huma γ3 hinge is encoded by three or four hinge exons (28). Analyses c genomic CH genes have led to speculations that the evolutionar history of H-chain genes may have included mutations that create or destroyed RNA splice sites and thereby converted portions c intron sequence into exon and vice versa. For example, th sequence of the intron 5′ to the hinge of the mouse γ2b gene show a surprising degree of similarity with the sequence of CH1; th observation led to the speculation (23) that the hinge exon ma have originated from a full Ig domain that became foreshortene either by the destruction of the RNA splice site at the 5′ end of th domain or the creation of a new splice site within the domain.

About 7 kb upstream from the murine Cμ gene lies a cluster c four JH segments (six JH segments in humans) that participate i VDJ recombination. Further upstream lie 13 D segments (about 2 in humans) and beyond them the VH regions. V and D regions ar described later in this chapter in the section on V-region diversity

Membrane Versus Secreted Immunoglobulin

Studies of IgH gene and cDNA structure have provided an expla nation for the alternative membrane and secreted forms of the I chain. As noted earlier, the membrane-bound forms of Ig H chain are slightly larger than the secreted forms owing to an additional C terminal hydrophobic segment that anchors the protein in mem brane lipids (29). In the case of the μ chain, these two forms ar products of two different mRNAs of 2.7 and 2.4 kb, which can b separated by gel electrophoresis. By comparing the DNA sequenc of a genomic μ clone and μ cDNA clones corresponding to thes two RNA species, several laboratories (30–33) have demonstrate that the two RNA species represent transcripts of the identical gen that have been spliced differently at their 3′ or C-terminal end (Fig. 6). The nucleotide sequence encoding the 20 C-termina residues of the secretory (μs) form is derived from DNA contigu ous with the CH4 domain of the μ gene, whereas in the membran mRNA (μm) the sequence after CH4 derives from two exons abou 2 kb 3′ further downstream. These membrane exons encode 4 residues, including a stretch of 26 uncharged residues that span th membrane to fix the Ig to the cell surface. The same general gen structure has been found for the other CH genes (34–37), sugges ing that the differential splicing mechanism probably accounts fo the two forms of Ig of all isotypes.

Early B cells make substantial quantities of both μm and μ whereas maturation to the plasma cell stage is associated with stron predominance of μs production, consistent with the function of suc cells in generating the pool of circulating Ig. The balance betwee the two RNA splice forms of μ has been interpreted as a competi tion between CH4-M1 splicing and the cleavage/polyadenylation a the upstream μs poly(A) addition site. The factors influencing thi balance have been studied by transfecting either early or late B cell

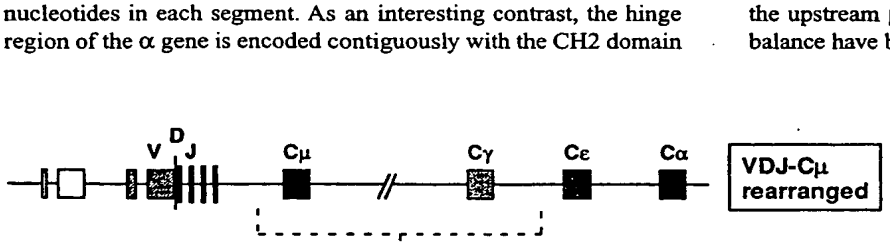


FIG. 5. Deletional isotype switch recombinatior The expression of downstream H-chain genes i accomplished by a recombination event tha replaces the Cμ gene with the appropriate H chain C gene (Cε as depicted here), deleting th DNA between the recombination breakpoints.

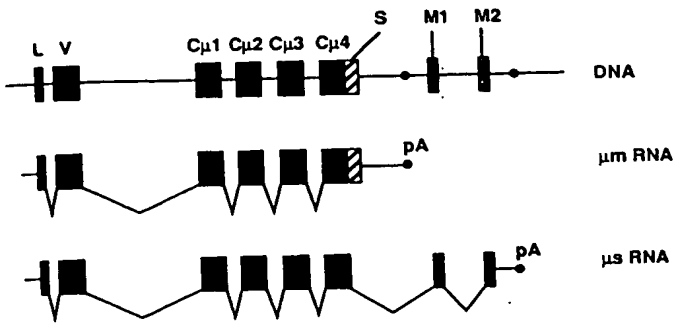


FIG. 6. Two RNAs generated from the μ gene by alternative processing. The top line illustrates the exons of the μ gene (*black rectangles*) in an expressed, rearranged μ gene. A primary transcript including all the exons present in the DNA can be processed as shown to yield either μs RNA (containing a C-terminal secreted [S] sequence) or μm RNA (containing the two membrane [M] exons).

with μ gene sequences or constructs in which the splice sites or cleavage/polyadenylation sites have been mutated, placed different distances apart, or rearranged in different order on the transfected gene construct. In some experiments constructs have been injected into frog oocytes with or without B-cell nuclei. Conflicting interpretations have emerged as to whether the critical factor influencing the μm/μs ratio is differential splicing (38) or poly(A) site choice (39,40). It does appear that the length of the intron between CH4 and M1 has an influence (41,42), that a stem-loop RNA structure at the 3′ end of the CH4-M 1 intron may play a role (43), and that the mechanisms regulating this ratio may be different for different isotypes (44). Additional investigations will be necessary to explain exactly how cell maturation leads to an appropriate alteration in the ratio of the membrane and secreted forms of Ig.

Membrane Ig serves as the antigen-specific component of the B-cell receptor (BCR), which is critical for initiating the signal for lymphocyte activation on contact with antigen, as described in Chapter 7. The segments of membrane Igs (of all isotypes) that penetrate into the cytoplasm are too short to encode functional signal transduction domains. Instead, transduction is mediated by an associated protein dimer composed of the BCR components Igα and Igβ. This dimer also has important signaling roles during B-cell development before the mature BCR is assembled, as discussed later in this chapter.

Organization of CH Gene Loci

As genomic clones for the C-region genes of the H-chain loci of humans and mice were obtained, efforts were made to "link" them, i.e., to clone continuous stretches of DNA including the CH genes as well as all the DNA lying between them in the genome. The general strategy of this work was to use cDNA clones to obtain the CH genes and to use "gene walking" techniques to fill in the noncoding DNA between the genes. The murine locus was completely linked in 1982 with a report (45) of clones covering the entire region of the mouse genome—all eight CH genes—spanning about 200 kb of DNA on chromosome 12. These clones define the general structure of the region as shown in Fig. 7, where the numbers indicate the distance in kilobases between the genes. All the CH genes are oriented in the same 5′ to 3′ direction. Recent sequence analysis has shown several γ pseudogenes within the clustered γ genes (46).

The human CH genes also have been cloned and localized to chromosome 14q32 (47), but not completely linked as of this writing. One significant difference between the human and murine IgH loci is that a large duplication exists in the human at the 3′ end of the H-chain gene locus, with two copies of a γ-γ-ε-α unit (48,49) (Fig. 7). One of the duplicated ε sequences is a pseudogene in which the CH1 and CH2 domains have been deleted (49,51). In addition, the human genome contains a third closely homologous ε-related sequence: a "processed" pseudogene found on chromosome 9 (50,51). (Pseudogenes of this type appear to have been reverse-transcribed from a processed RNA intermediate and then to have been inserted in the genome at locations unrelated to the original locus of the transcribed source gene.) A γ-related pseudogene lacking a switch region is also present in the human IgH locus between the two γ-γ-ε-α duplications (52). The map presented in Fig. 7 is based on partial contiguous overlaps and pulsed field gel

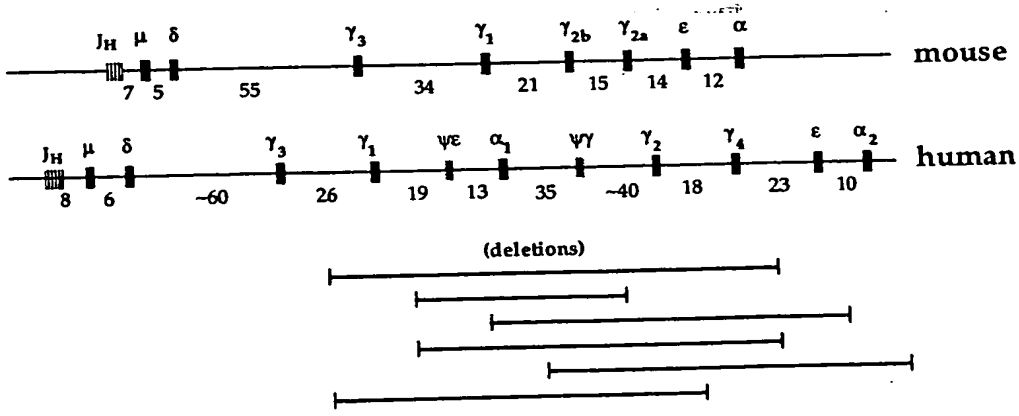


FIG. 7. H-chain constant region loci of mice and humans. The murine locus has been cloned in its entirety (45); the constant region genes are diagrammed with the approximate intergene distance indicated below (in kb). The human locus shows a large duplication at the downstream end. Although not all of the diagrammed segments have been physically linked by contiguous clones, the indicated order is supported by the deletions observed in various individuals as well as mapping by PFGE (53,54).

electrophoresis (PFGE) (53,54). PFGE allows electrophoretic separation of very large fragments—up to several megabases in length—that cannot be separated by conventional electrophoresis; fragments of this length are useful in mapping over long distances and can be generated by restriction enzymes that have unusually rare recognition sequences. The map of the human IgH locus in Fig. 7 is consistent with the known deletions in the H-chain locus (55), as diagrammed in the same figure.

The IgH locus also has been examined in several other species besides mice and humans, and several notable differences have been observed. Rabbits, for example, have 13 Cα sequences and only a single Cγ gene (56); this unusual expansion of genes contributing to mucosal immunity may be related to the peculiar habit of coprophagy in these animals. In contrast to the multiplicity of rabbit Cα genes, pigs have only one Cα gene and eight Cγ genes (57). Camels are unusual in having H chains that function in the absence of L chains (58). H-chain Ig genes (VH or CH) have been cloned from a number of other species, including rats [which are highly homologous with mice (59)], cows (56), chickens (60–63), horses (64), sharks (65), bony fish (66,67), crocodiles (68), frogs (69) and axolotls (70).

Heavy-Chain Switch

Switch Regions

The availability of genomic IgH clones allowed detailed sequence analysis of the deletional switch recombination. The active switched genes from several myelomas were compared with the corresponding germline CH genes and with the germline μ gene, with particular attention to the sequences surrounding the switch recombination site. In each case the recombination events were found to have occurred within or near regions of remarkably internally repetitive DNA sequences 5′ to the CH coding sequences; these have become known as switch (S) sequences (71–74).

The S region of the mouse μ gene, Sμ, is located about 1 to 2 kb 5′ to the Cμ coding sequence and is composed of numerous tandem repeats of sequences of the form $(GAGCT)_n(GGGGT)$, where n is usually 2 to 5 but can be as high as 17 (74). These repeats apparently promote deletions within the Sμ region by homologous recombination events that can occur during the laboratory construction and isolation of clones containing the Sμ region. Because of such deletions, most cloned germline μ genes are found on *Eco*RI fragments shorter than the 12.5-kb fragment identified in genomic blots of BALB/c DNA. Deletions of the same region have been demonstrated to occur in vivo by comparison of the μ locus in different mouse strains by Southern blotting (75) and may occur especially frequently during the activity of switch recombination in normal B cells (76).

Similar internally repetitive S regions spanning 1 to 10 kb have been found 5′ to all the other CH genes except Cδ. All of the S regions include occurrences of pentamers similar to GAGCT and GGGGT that are the basic repeated elements of the Sμ gene (77); in the other S regions these pentamers are not precisely tandemly repeated as in Sμ, but instead are embedded in larger repeat units. The 10-kb Sγ1 region has an additional higher order structure: two direct repeat sequences flank each of two clusters of 49-bp tandem repeats (78). S regions of human H-chain genes have been found very similar to their mouse homologs (79–81). Indeed, sequence similarity between human and mouse clones 5′ to the CH genes has been found to be confined to the S regions, an observation that supports the biologic significance of these regions.

A switch recombination between, for example, μ and ε genes produces a composite Sμ-Sε sequence (Fig. 8). By examination of the germline Sμ and Sε sequences in comparison with the myeloma- or hybridoma-derived Sμ-Sε composite S region, it has been possible to localize the exact recombination sites between Sμ and Sε that occurred in different cells; similar analyses have been performed with cells producing other isotypes. These studies have indicated that there is no specific site, either in Sμ or in any other

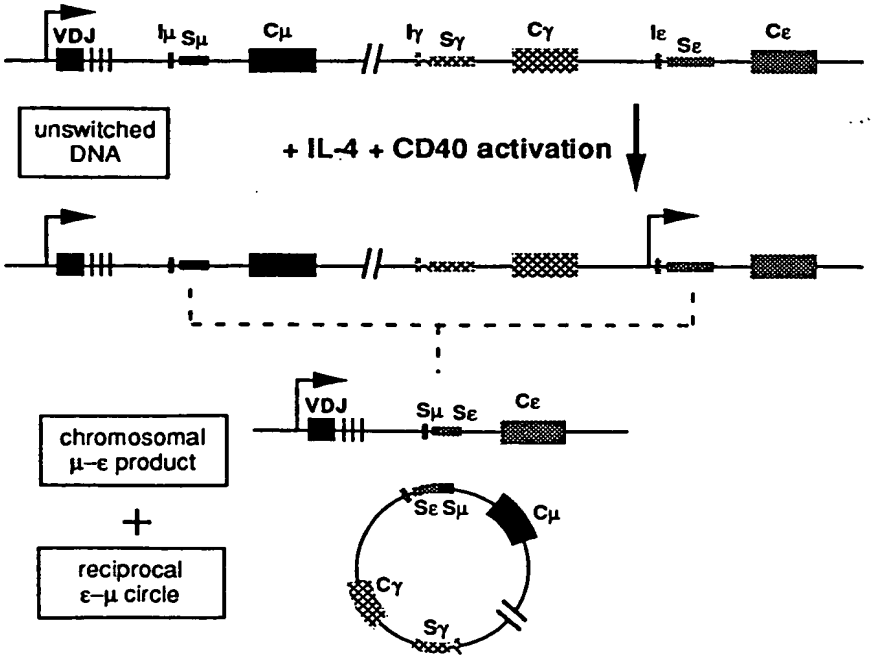


FIG. 8. Switch regions and composite switch junctions. The recombination breakpoints in isotype switch recombination fall within repetitive S regions. Stimuli that activate switch recombination (IL-4 and CD40 activation in the example shown) generally promote transcription across the target S region, initiating just upstream at the I exon. Recombination between Sμ and Sε produces two composite switch junctions: an Sμ-Sε junction retained in chromosomal DNA, and a reciprocal Sε-Sμ junction found in fractions of circular DNA. PCR amplification across either composite junction can be used to study switch recombination.

S region, where the recombination always occurs. Thus, unlike the enzymatic machinery of VJ recombination, the switch machinery can join sequences in a broad target region; this makes sense because VDJ recombination occurs within coding sequences, whereas switch recombination is less constrained because it occurs in introns. Many composite switch junction sequences show evidence of mutations at the recombination breakpoint when compared with the corresponding germline switch sequences; these mutations have been interpreted as reflecting an error-prone DNA synthesis step that may be a component of the switch recombination mechanism (82).

DNA excised by switch recombination has been detected by cloning from fractions of circular DNA isolated from cells actively undergoing isotype switch recombination (83–85). Thus, at least some of the excised DNA segments ligate their ends to form switch circles; these contain composite switch junctions that are reciprocal to the composite switch junction retained on chromosomal DNA (Fig. 8). For the example of cells switching from μ to ε, composite Sμ-Sε junctions are found on chromosomal DNA, whereas Sε-Sμ junctions can be found representing the reciprocal junctions from switch circles. Because switch circles are not linked to centromeres and may not contain origins of replication, they are not efficiently replicated. Therefore, they are not found in cells that have divided multiple times after switching, e.g., in myelomas or hybridomas.

Methods of Assaying Switching

In stable myelomas or hybridomas expressing switched isotypes, evidence of switch recombination can be obtained by gene cloning or Southern blotting. However, for studies of the regulation and mechanism of switch recombination, assays are needed that can detect switch recombination in a minority population of cells switching in culture. Some laboratories assess switching by simply measuring Ig protein of the switched isotype appearing in the culture supernatant. Alternatively, reverse-transcriptase polymerase chain reaction (RT-PCR) can be used to detect mRNA corresponding to the mature VDJ-C RNA transcripts of the switched isotype. However, because the culture conditions favoring isotype switching also may influence transcription or protein synthesis rates independently of switch recombination, RNA or protein assays may not faithfully reflect the DNA recombination events. Furthermore, switched RNA or protein cannot be assumed to reflect DNA recombination if one is exploring one of several models for nonrecombinational mechanisms for isotype switching. Therefore, two different PCR strategies have been developed to assess switch recombination at the DNA level. In one strategy, PCR primers are designed to amplify across the composite S region of interest (86). A related strategy is to amplify the reciprocal switch junctions found on circular DNA (87,88); these junctions can be used to "count" recombination events independent of proliferation if one assumes that each circle is produced as a by-product of a single switch recombination event and, failing to replicate as the cells divide, is randomly partitioned to daughter cells at successive divisions after the recombination event. Because the efficiency of amplification varies for different composite switch junctions—smaller templates are amplified more efficiently, and the largest may not amplify at all—the PCR strategy described above cannot easily be adapted to assay switch recombination quantitatively. For this reason a second strategy known as digestion-circularization PCR (DC-PCR) was developed (89). In this approach DNA from switching cells is digested with a restriction enzyme, and restriction fragments—including the ones bearing a composite Sμ-Sε junction—are ligated to form circles. Primers designed to amplify across the restriction site generated by ligation of the Sμ-Sε fragment ends will yield a consistent product whose size depends only on the distance between primers and the restriction site. From unswitched DNA no product is amplified because the two primers can never both hybridize to the same DNA circle. Therefore, with appropriate calibration (90), the amount of DC-PCR product formed can be used as a semiquantitative measure of the amount of composite switch junctions in a DNA sample. These methods have been used in many of the experiments described below.

Regulation of Isotype Switching

Isotype switching occurs physiologically in animals about 1 week after immunization with T-dependent antigens, at about the same time that somatic mutation of Ig genes begins. Somatic mutation (discussed later in this chapter) clearly occurs in germinal centers of lymphoid organs—a location that facilitates T- and B-cell interaction—and there is some evidence that germinal centers are a major site for isotype switching as well. As demonstrated by in vitro switching experiments, T cells promote switching by secretion of cytokines (especially interleukin [IL]-4 and transforming growth factor-β [TGF-β]) as well as by cell-to-cell contact. A major component of the cell contact signal is mediated by an interaction between the B-cell surface marker CD40 and its ligand (designated CD40L or glycoprotein [gp]39), expressed on activated T cells. The dependence of switching on this interaction is highlighted by the genetic disease known as the X-linked hyper-IgM syndrome, which was found (independently by several laboratories) to be caused by a defect in the gene encoding CD40L/gp39 (91). Patients with this syndrome have elevated concentrations of IgM in their serum and almost no Igs of other isotypes. In addition, their antibodies fail to show affinity maturation or evidence of B-cell memory responses. Mouse strains with engineered defects in CD40 or CD40L show a similar phenotype, although they respond with normal isotype switching to T-independent antigens (92); little is known about this T-independent switching pathway. The discovery of the importance of the of CD40–CD40L interaction has facilitated in vitro switching experiments in which T cells can be replaced with antibodies to CD40 or with cells engineered to express surface CD40L. One role of the CD40 stimulus is to induce B-cell proliferation. Indeed, other proliferative stimuli (e.g., lipopolysaccharide [LPS] or IgM or IgD cross-linking) can support cytokine-induced isotype switching in vitro in the absence of T cells and CD40 activation; and switching may be related to the cell cycle (93). However, CD40 has additional effects, including upregulation of IL-4 responsiveness and IL-4 receptor number (94); the signaling pathways initiated by CD40 are under active investigation (95).

Different isotypes are known to predominate in different immune responses depending on the antigen, route of antigen administration and several other parameters. As discussed more fully in Chapter 23, these different parameters act in part by influencing the cytokine milieu of the B cells. IL-4, for example, promotes the expression of IgE (and IgG1 in mice), whereas TGF-β promotes switching to IgA. These lymphokines are believed to act by making the C region of the target isotype accessible to switch recombinase machinery that may be non–isotype-specific. The accessibility is associated with expression of an RNA transcript that initiates upstream of a target S region and extends through the target C region (Fig. 8). This type of RNA is designated a germline transcript because it is transcribed while the

IgH locus is in germline (i.e., unswitched) configuration; alternatively, these transcripts are called sterile (i.e., lacking a V region). After in vitro treatment of B cells with IL-4, for example, but before any switch recombination to Cε, sterile transcripts are detected with a structure that includes Cε preceded by a short exon known as Iε. The Iε sequence derives from DNA upstream of Sε, a location that would be deleted during the formation of the Sμ-Sε composite S region; in the germline transcripts the Iε region is spliced to Cε by removal of an intron containing the Sε region. Similar transcripts have been found for every isotype examined in both human and mouse systems, including μ. In each case many of the same experimental conditions (including cytokines) that favor the accumulation of sterile transcripts from a particular isotype also favor switch recombination involving the corresponding S region. In some cases the signals transduced by the cytokine receptor have been elucidated. For example, IL-4 stimulates transcription by activating the transcription factor STAT6, which attaches to one of several nuclear protein binding motifs in the promoter region upstream of Iε and Igλγ, as discussed later in this chapter. Apart from I-region promoters, sterile transcription and isotype switching are also regulated by an enhancer lying downstream of the murine Cα gene, as deduced from switching defects in mice in which this enhancer was replaced by a neomycin resistance gene in all B cells (96); defects were observed in switching to IgE and several IgG isotypes, but not to IgG1.

Studies of mouse strains in which the I region from various isotypes have been targeted by homologous recombination suggest that sterile transcription is necessary but not sufficient for recombination (97–99). The low extent of sequence conservation of the I exons and the lack of consistent open reading frames suggest that these transcripts do not encode a functional protein. What then is their role? One hypothesis is that the critical chromosomal alteration that renders an isotype locus accessible to the switch recombinase machinery is achieved by the process of transcribing through the locus, and that the transcripts themselves serve no function. A second hypothesis is that the transcripts participate in the recombination event in some way, perhaps by formation of an RNA:DNA triple helix (100). In support of this idea, cell-free transcription of S regions was found to lead to a stable association of the transcript RNA with the template DNA (101); significantly, this association occurred only with RNA transcribed from S region DNA and only when the RNA was transcribed in the physiologic orientation. Neither of these two hypotheses concerning the role of sterile transcription account well for a feature conserved in all the transcripts: the RNA splice that removes the Sx region from the mature IxCx transcript. It is noteworthy that sterile transcripts from the germline components of V(D)J assembly recombination are also synthesized just before that recombination event, and transcription is also observed from rearranging yeast DNA sequences. These observations suggest that the transcription of DNA immediately before recombination may be a general feature of recombination events common to many biologic systems. On the other hand, it is likely that cytokines regulate other aspects of the switching mechanism besides sterile transcription because several examples have been reported of cytokines up- or downregulating switch recombination without a parallel effect on sterile transcripts (102).

Mechanism of Switch Recombination

The mechanism of isotype switch recombination has been probed with a variety of strategies, so far with limited results. One approach to delineating the sequences required for switch rearrangement has been to construct plasmid substrates containing switch sequences that might undergo switch recombination when transfected into B-lineage cells either stably (103) or in transient systems (104). For example, the construct of Daniels and Lieber (105) contained the polyoma origin and T-antigen gene (to allow replication in murine cells) and fragments of Sμ and Sγ3 segments, with viral promoters upstream of each and a *supF* transfer RNA (tRNA) gene between them; expression of the *supF* tRNA gene in appropriately engineered bacteria led to blue colonies on culture plates. Plasmids undergoing Sμ-Sγ3 recombination in eukaryotic cell lines could be recovered and identified by the production of white colonies in bacteria. Although various nonlymphoid cell lines produced white colonies within the first 20 hours of transfection (perhaps resulting from DNA repair enzymes acting on nicked plasmids), continued increases in the percentage of white colonies beyond 20 hours appeared to be B cell specific. Deletion of the promoters had only minor effects on the recombination frequency, but white colonies were dramatically decreased when the promoters were arranged so that the S regions were transcribed in the nonphysiological direction. When the S regions were replaced with irrelevant DNA, the direction of transcription had no effect on recombination. The dependence of recombination frequency on transcriptional orientation of switch sequences parallels findings described above in which the RNA–DNA complex involving S regions was strand dependent.

Another strategy for elucidating the switch recombinase mechanism is to identify intermediates in the reaction, an approach that has been strikingly successful in studying VDJ recombination, as discussed later in this chapter. A single study exploring this approach has used ligation-mediated PCR (a technique described later in connection with VDJ recombination) to detect blunt, double-stranded cuts in the murine γ3 region in B cells switching in culture (106); these cuts may be generated by the switch recombinase machinery.

Possible Switch Recombinase Components

In an effort to identify components of the recombinase machinery, several laboratories have investigated proteins that bind in a sequence-specific manner to S-region sequences in vitro. Several examples are described below, although it should be emphasized that none of the components discussed in this section has been demonstrated to participate in switch recombination. LR1 is a protein found in nuclear extracts from murine splenic B-lymphocytes after induction with LPS; it binds to Sγ1, Sγ3, and Sα, as well as to the H-chain enhancer (107). The protein has been purified (108), and one component has been identified as the nucleolar protein nucleolin (109). Sμbp-2 is a ubiquitous protein, also upregulated by LPS in murine splenic B cells, which binds to a segment from the tandem repeats in Sμ. A murine cDNA clone was found to exhibit sequence similarity to genes encoding helicases; such an activity could be critical for switch recombination (110). NF-Sμ is another protein that binds to Sμ tandem repeats and is induced in splenic B cells by LPS; its binding specificity is slightly different from the other proteins already described (111). Two proteins that bind within Sγ regions to subsequences associated with a high frequency of recombination junctions have been designated SNIP and SNAP and apparently correspond (respectively) to the transcription factors NF-κB/p50 and E47, which are discussed later in this chapter (112,113). A role for NF-κB in switching is supported by experiments in B cells from a mouse strain in which the p50 gene has been disrupted by homologous recombination. In these p50 knockout mice, isotype switching to IgE and IgG3 secretion was

markedly reduced; however, reduced expression of the corresponding germline transcripts could indicate that the p50 was required for promoting "accessibility" rather than for the actual recombination event (114). In these experiments switching to IgG1 expression was almost unaffected by the absence of p50, and α was the only isotype whose expression was markedly reduced in the face of normal germline transcription. A possible role for E47 in switching is supported by experiments in which expression of Id1, an antagonist of the E2A transcription factors of which E47 is a member, was found to partially inhibit spontaneous and induced switching to IgA in the murine cell line CH12.LX2 (115).

Apart from studying proteins that bind to S-region DNA, another approach to identify switch recombinase components has been to search for an enzyme activity expected to participate in the recombination. A lymphoid-specific endonuclease activity that preferentially cleaves G-rich segments of S regions has been partially purified and proposed as a possible participant in switch recombination (116).

The possibility that switch recombination depends on some of the same components that are known to participate in V(D)J recombination has been tested for several proteins (whose role in V(D)J recombination is discussed later in this chapter). Both SCID mice, which are natural mutants of DNA-dependent protein kinase (DNA-PK), and mice with homozygous knockouts of their recombinaton activating gene-2 (RAG-2) genes are impaired in developing mature B lymphocytes because of their inability to assemble V genes efficiently. However, when early B-lineage cells from these mice were allowed to proliferate *in vitro* and were then treated with IL-4 and anti-CD40, switch recombination occurred in the RAG-2 knockout cells but not in the SCID cells (116a). Thus, DNA-PK appears necessary for switch recombination but RAG-2 does not. DNA-PK binds DNA as part of a complex that also contains the protein Ku80 (also discussed later in this chapter). Recently, Ku80 was also implicated in switch recombination in experiments in which Ku80 knockout mice were crossed with mice in which recombined VκJκ and VDJ genes were "knocked in" to the respective loci by homologous recombination. Whereas "knock-in" mice with intact Ku80 genes expressed IgM encoded by the engineered Vk and VH genes and also switched to downstream isotypes, the corresponding Ku80-deficient mice made IgM but did not switch isotypes, suggesting that Ku80 is also required for switch recombination (116b).

A recent achievement that holds promise for identifying recombinase proteins is the development of a cell-free nuclear extract system that can accomplish recombination between S-region sequences in vitro (117,118). This system depends on a powerful assay in which tritium-labeled plasmid molecules containing Sγ are incubated with digoxigenin-labeled plasmid containing Sμ. Recombination between the plasmids is detected as tritium immunoprecipitable by antidigoxigenin, and the recombinant DNA structure can be verified by PCR amplification across the composite switch junctions. Optimal recombination was found to require adenosine triphosphate (ATP), both Sμ and Sγ, and nuclear extract from LPS-blasted B cells. Partial fractionation of nuclear extracts identified an active complex designated SWAP (switch activation proteins) composed of at least four proteins: nucleophosmin (which has a RecA-like DNA D-loop forming activity), poly(ADP)ribose polymerase (PARP, a nuclear protein implicated in DNA repair), nucleolin (described above as a component of the Sμ-binding protein LR1), and a novel 70 kD protein designated SWAP-70, not homologous to any known protein family (118). SWAP-70 is strongly expressed only in B cells that have been activated for switch recombination, and binds with high affinity to the other components of the complex. The SWAP complex is a strong candidate for a switch recombinase component, but as of this writing, the definitive evidence from a knockout experiment is not yet available.

Nonstandard Switch Recombination

Thus far we have considered switch recombination to involve a simple deletion of the DNA between two S regions; although this is the most common scenario, three additional situations should be considered for completeness.

Sequential Switching. Several switch recombination events can occur sequentially on a given chromosome. One well-studied example involves sequential switching to γ1 followed by ε in mouse B-lymphocytes. The same cytokine, IL-4, promotes switching to both isotypes. After an initial switch recombination generating a composite Sμ-Sγ1 junction (leading to IgG1 expression), this composite S region can undergo a secondary switch recombination with Sε, which lies downstream. In IgE-expressing cells, evidence of the initial recombination to γ1 can be demonstrated by the presence of a composite Sμ-Sγ1-Sε junction (119), or by the detection of the reciprocal switch circle product Sε-Sγ1. To assess the quantitative importance of this pathway in IgE generation, resting B cells stimulated with IL-4 plus LPS were treated with an anti-IgG1 antibody to eliminate cells expressing this isotype from the culture; IgE secretion was inhibited about 70%, suggesting that most mouse B cells expressing IgE have undergone an intermediate stage in which they expressed IgG1 (120). However, in mutant mice with a block in γ1 switching due to a targeted deletion in the γ1 locus, the frequency of switching to ε is normal, suggesting that the sequential switching results from the simultaneous accessibility of both Sγ1 and Sε, rather than an obligatory sequential switch program (121). Sequential switching to IgE expression via IgG also occurs in human B cells (81,122), but the quantitative significance of this pathway is not known.

Inversional Recombination. Some switch recombinations apparently lead to inversion rather than deletion of the DNA between the two S regions involved (123,124). A chromosome with an inversional switch recombination would be incapable of encoding a functional H chain because the C region downstream of the VDJ region would be in inverted orientation, but the chromosome could be "rescued" by a second switch recombination to a downstream constant region. In human B-cell leukemias—which are under no selection for Ig production—inversional switch recombination has reported to occur at a frequency of about 15% (125).

Trans-Switching. Although most switch recombinations involve a single chromosome, transchromosomal switching between allelic chromosomes has been detected in rabbits at a frequency of about 5% (126). The detection of trans-switching in rabbits was facilitated by the availability of allotypic markers of C and V regions in this species; the frequency of trans-switching in other species is not known.

Switched Isotypes Without Switch Recombination

Several laboratories have reported detection of B cells expressing Ig of more than one isotype. Such double-producing cells may reflect a normal transient intermediate stage when a switched isotype may be expressed (after normal switch recombination) along with IgM that is retained in the cells because of the long half-life of the protein or its mRNA. However, some laboratories have reported a stable double-producer phenotype in cell lines without apparent switch recombination in the expressed IgH locus. For the

case of μ-δ double producers, the explanation is apparently that δ transcripts can be produced by RNA splicing from a long primary transcript that includes μ and δ (127). More difficult to explain are the cell lines expressing μ along with an isotype whose C region is so distant from Cμ that an analogous long transcript would be on the order of 100 kb or more; such transcripts are longer than have been observed with current laboratory methods, although precedents for genes whose exons are spread over similar distances are known. One interesting proposal being considered to explain expression of downstream isotypes without switch recombination is that separate short transcripts of VDJ and a downstream CH gene (i.e., a sterile transcript) could be joined by a trans-splicing mechanism similar to that documented for trypanosomes and certain viruses (128,129). A nonphysiologic mechanism has been described that could account for some cases of double isotype production as a consequence of chromosomal duplication (130). Stable double-producing cell lines continue to be studied (131–133); at present we cannot be certain whether cells stably expressing this phenotype represent important physiologic counterparts of normal B cell subsets. A semiquantitative assessment of switch recombination in a population of murine B-cell switching in vitro to IgG1 indicated that DNA rearrangement can account for the IgG1 expression observed (90), suggesting that most expression of switched isotype Ig is associated with switch recombination, and that alternative nonrecombinational models for switching do not seem to be required on quantitative grounds.

Kappa Light-Chain Genes

In comparison with the H-chain genes, the κ locus is relatively simple. A single Cκ gene with a single exon and no reported alternative splice products is found in both mice and humans. Upstream of the murine Cκ gene lie five Jκ gene segments, spaced about 0.3 kb apart (5,6). Of these Jκ segments, the third encodes an amino acid sequence never observed in κ chains and is believed to be nonfunctional owing to a defect in the splice donor site that would join the corresponding RNA sequence to Cκ. The human locus is similar, with five Jκ regions upstream of Cκ; however, no homolog of the defective murine Jκ3 is present in the human Jκ cluster, whereas an additional Jκ sequence lies downstream of the sequence homologous to murine Jκ5 (134,135). Upstream of the Jκ segments in both species lie the Vκ genes, which will be described later in this chapter.

Apart from Vκ-Jκ rearrangement, an additional recombination event occurs in this locus, a recombination unique to κ genes and apparently mediated by the same 7-mer/9-mer signal elements involved in V(D)J recombination. This event, which involves deletion of the Cκ gene segment, was initially suggested by the observation that Southern blots of DNA from λ-expressing human lymphoid cells generally show no detectable Cκ sequence (136). Apparently in most B cells the Cκ genes are deleted from both chromosomes before λ gene rearrangement begins. When the boundaries of the deleted segment of DNA were examined in several human and mouse cell lines, a common sequence element was found at the downstream boundary; this element was designated RS (recombining sequence) in the mouse studies (137) and κde (kappa-deleting element) in the human studies (138). The human κde in germline DNA is located 24 kb downstream from the Cκ gene and is flanked by a 7-mer/9-mer RSS similar to that found flanking the Jκ regions (i.e., with a 23-bp spacer) (139). The similar murine RS is about 25 kb downstream from murine Cκ (140). The κde element can apparently recombine either with a Vκ gene segment (leading to a deletion of the entire Jκ-Cκ locus) or with an isolated 7-mer element that is located in the Jκ-Cκ intron (leading to deletion of Cκ but retention of the Jκ locus). The 7-mer in the Jκ-Cκ intron is 30 bp 5′ from a poorly conserved 9-mer-like sequence, a spacing that seems to violate the usual 12/23 rule. The significance of this unusual spacer is not understood, but possibly the 7-mer in these recombinations is active without a functional 9-mer, as seems to be the case for secondary VH recombinations (discussed in a later section).

A comparison between the mouse RS and human κde sequences (138,139) shows that the recombination signals are highly conserved and that downstream of these elements a region of about 500 bp is partially conserved (about 50% sequence identity). The latter region includes open reading frames of 127 (mouse) or 102 (human) codons. It is not known whether these reading frames are ever expressed as protein as a consequence of the RS/κde recombination events, but the fact that the recombination may occur with either a Vκ region or intron sequence suggests that the sequences joined by the event may be less important than the sequences deleted. RS/κde elements are consistently found to be rearranged in cells in which Cκ is deleted and λ rearrangements are found; this has led to the speculation that the RS/κde recombination event may mediate the developmental switch from κ to λ gene rearrangement, perhaps by deleting a gene for a negative regulator of λ gene rearrangement. However, current evidence argues against this view.

Lambda Light-Chain Genes

Murine λ Locus

In laboratory mouse strains, λ chains represent only about 5% of L chains, and this diminished abundance is associated with remarkably meager diversity. In contrast to the κ system with its multiple V-region families, amino acid sequence analysis of monoclonal λ chains detected only two sequences that appeared to represent germline Vλ regions. Furthermore, in contrast to the single mouse Cκ region, three nonallelic mouse isotypes are known from secreted λ chains; these are designated λ1, λ2, and λ3, in decreasing order of abundance.

The first λ gene to be cloned was a germline Vλ2 gene obtained by Tonegawa's laboratory in 1977 (141) (this was the first Ig gene cloned). The sequence of this Vλ gene (7) showed structural features that are similar to those of other germline Vλ genes, as well as Vκ and VH genes, which were discovered later. The Vλ2 coding sequence begins with a 19-amino-acid signal peptide that is interrupted within codon 4 by an intron (which was one of the first introns demonstrated). After the remaining signal peptide codons, the DNA sequence matches closely that expected based on amino acid sequence determined chemically from a λ2 myeloma L chain. However, the sequence of this germline Vλ2 gene ends abruptly 13 codons short of the expected end of the Vλ2 region, an observation that led to the first recognition of a separately encoded J region.

Cloning and long range mapping studies by pulsed field gel electrophoresis (142,143) have led to a substantial understanding of the mouse λ locus (Fig. 9). There are four Cλ genes, each with its own Jλ-region gene located about 1.3 kb 5′ from the C. The J-Cλ3 and J-Cλ1 genes are arranged in one cluster about 3 kb apart with the Vλ1 gene lying about 19 kb upstream. A second Cλ cluster lying about 130 kb upstream from the Cλ3-1 locus contains J-Cλ2 and an unexpressed gene J-Cλ4. These are flanked by two upstream Vλ genes, Vλ2 and the rarely used Vx, which has an in-frame termination codon at its 3′ end (144). The gene order (V2-Vx-JC2-JC4-V1-

JC3-JC1) explains the common expression of Vλ2 (or Vx) in association with Cλ2 and Vλ1 with Cλ1 or Cλ3. The Vλ2 has been found in rare association with the 190-kb distant Cλ1 locus, but the backward recombination of Vλ1 with Cλ2 has not been observed. The similarities between the four J-C genes suggest that the two clusters arose by a duplication of an ancestral V-J-Cλx-J-Cλy unit that in turn was the result of a prior J-Cλ duplication event. The ancestry of the Vx gene is uncertain because this gene is rather dissimilar to the other Vλ genes; indeed, it resembles Vκ as much as Vλ. Anti-Vλx antisera detect expression of this Vλ in all laboratory mice tested, but it may have a particular restricted function.

The sequences of genes in the λ locus have been examined for clues that might explain the relative abundance of their expressed products—λ1 > λ2 λ3 > > > (λ4) (145). The sequence of Cλ4 includes several amino acid substitutions but no termination codons that would necessarily render it nonfunctional; however, at the 3′ end of Jλ4, a mutation has destroyed the "GT . . ." found at almost all known donor splice sites, so that an RNA transcript of this gene would not be properly processed (reminiscent of the mouse Jκ3). The Jλ gene segments are all flanked on their 5′ sides by sequences similar to the 9-mer and 7-mer signal elements observed in the VH and Vκ system. The 12/23 rule discussed in relation to spacing between the signal elements in the κ locus also applies to λ, but in the λ locus the RSS elements are spaced about 23 bp apart for V regions and about 12 bp apart for J regions (the opposite of the arrangement in κ genes), as shown in Fig. 4. The decreased abundance of λ2 and λ3 relative to λ1 may be related to discrepancies between their 9-mer homology elements and the consensus 9-mer element.

Analyses of λ genes in wild mice by Southern blotting have indicated more complex and varied loci than that seen in typical laboratory strains (146). These complex λ loci may result from gene duplication events beyond those evident in laboratory strains, although the observation that at least one wild Vλ gene missing from BALB/c is similar to a human Vλ (146) suggests that some of the difference between wild and laboratory strains must be due to gene loss in that latter.

Human λ Locus

Lambda L chains are much more abundant in humans than in mice (about 40% of human L chains versus about 5% in mice). Furthermore, four isotypic forms of human λ chains have been characterized, known by their original serologic designation as Kern–Oz–, Kern–Oz+, Kern+Oz–, and Mcg; several other variants have been described, perhaps representing allelic polymorphisms.

Seven human Jλ-Cλ segments are clustered within an approximately 33-kb region of DNA that has been entirely sequenced (147–149). As shown in Fig. 9, genes for the four major expressed human λ isotypes have been localized within the major cluster and correspond to JCλ1, JCλ2, JCλ3, and JCλ7. The remaining three homologous J-C segments are apparently pseudogenes, with either in-frame stop codons or frame-shifting deletions. However, JCλ6 may be functional in some individuals (150), and the common allele—which has a 4-bp insertion leading to a deletion of the C-terminal third of the Cλ region—can nevertheless undergo Vλ-Jλ recombination, encoding a truncated protein that can associate with H chains (151). A variety of polymorphic variants of the human λ locus have been detected, apparently the result of gene duplication; as shown in Fig. 9, one to three extra λ segments have been detected on Southern blots of human DNA (152).

Three Cλ-related sequences have been discovered near the major Jλ-Cλ cluster. One of these, designated λ14.1, represents the human homolog of the murine surrogate L chain λ5. Finally, an additional weakly hybridizing DNA segment outside the linked cluster has been characterized as a processed pseudogene (153). V genes of the human λ system have been completely characterized, as discussed in a later section.

λ-Related Surrogate Light Chains

Immunoglobulin μ H chains can be detected on the surface of pre-B cells that do not make L chains. However, in mature B cells,

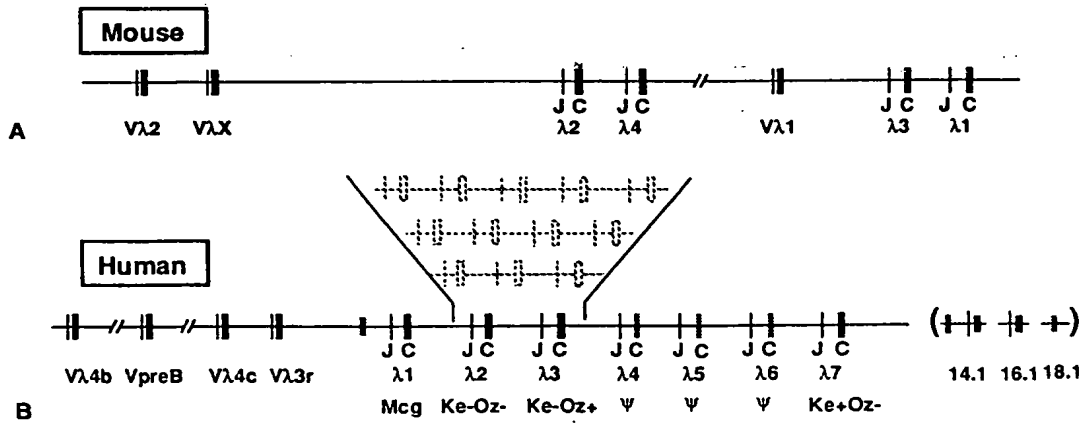


FIG. 9. Germline λ genes. The maps in this figure are schematic, i.e., not to scale. **A:** The murine λ gene system includes four JC complexes and three V genes, which have been characterized in two unlinked contig (sets of overlapping clones) as shown. **B:** The human λ locus has been characterized by complete sequence analysis. The human VpreB surrogate L-chain gene is located within the Vλ cluster. The Cλ locus includes a segment of seven JC complexes plus three additional unlinked sequences. The hatched JC complexes diagrammed above the seven linked λ sequences represent polymorphic variants with additional duplications of the JC unit as deduced from Southern blots. The 14.1 sequence—the human λ5 surrogate L-chain homolog—lies downstream from the JC cluster but its location relative to the other I-like sequences is not known. Exon 1 of the 14.1 gene is homologous to a exon upstream from Jλ1 (*gray rectangle*).

Ig H chains cannot reach the cell surface if L-chain synthesis is interrupted. What allows the expression of surface H chain in pre-B cells in the absence of L chain?

The first clues to this question were uncovered in a search for genes whose expression is specific to the pre-B stage of lymphocyte development. Melchers and colleagues identified one such gene that demonstrated striking sequence similarity to the J and C regions of the λ locus; they named it λ5 because four murine Cλ genes were already known (154,155). The genomic λ5 gene includes three exons (Fig. 10): exon 1, which appears to encode a signal peptide; exon 2, whose 3′ end is homologous to Jλ; and exon3, homologous to Cλ. When flanking regions of the genomic λ5 clone were tested as probes against pre-B cell mRNA, another transcribed segment was found about 4.7 kb 5′ from λ5 (156). (Fig. 10.) Sequence analysis of the latter region showed similarities to both Vλ and Vκ; for this reason (and because of its expression in pre-B cells) it is called VpreB1. A second, nearly identical sequence in the mouse genome is named VpreB2 (157) and appears to be functional (158); a less similar VpreB3 also has been described (159). Neither λ5 nor VpreB genes show evidence of gene rearrangement in B or pre-B cells. Both genes have typical consensus splice sites and initiation and termination codons and have no apparent defects that would prevent their expression as proteins. That they are expressed and serve an important role is suggested by the conservation of homologs in every mammalian species examined.

Evidence strongly supports the notion that these genes encode surrogate L chains (SLCs) that associate with μ H chains to permit surface μ expression before the availability of L chains. Thus, when a μ H-chain gene was transfected into an Ig-negative myeloma line, no surface μ expression was observed unless λ5 and VpreB genes were also transfected (160). The surface μ chains were found to be covalently linked to the 22-kDa product of the λ5 gene, whereas the 16-kDa VpreB product was noncovalently associated. A similar complex is observed in pre-B cell lines and in normal bone marrow pre-B cells (161). The V-like VpreB gene product [also known as ι (162)] apparently associates with the Cλ-like λ5 product (also known as ω) to form an L chain-like heterodimer that can fulfill some functions of a true L chain.

One likely role for a μ-SLC complex is suggested by the observation that most Vκ-Jκ recombination occurs only in cells expressing a functional μ H chain (as discussed more fully in the section on regulation of V(D)J recombination); apparently μ-SLC expression on the cell surface can trigger the onset of Vκ-Jκ rearrangement. Evidence for this view comes from experiments in which pre-B line that normally does not rearrange its κ locus was transfected with a construct encoding the membrane form of μ H chain (163); when the transfected μ gene was expressed in a complex containing VpreB and λ5, Vκ rearrangement was induced. In contrast, surface expression of a deleted μ gene (μΔm)—which lacked VH and CH1 and which did not associate with SLC—was ineffective in inducing Vκ rearrangement unless the μΔm protein was crosslinked by an anti-μ antibody. These results suggest that the SLC may facilitate cross-linking of surface μ chains in pre-B cells, a necessary step before the B cell can proceed to Vκ rearrangement, L-chain synthesis and mature Ig production. Further support for such a critical role is discussed later in this chapter.

Human homologs of both λ5 and VpreB have been cloned. Three λ5-like sequences are located downstream from the Cλ cluster on chromosome 22 (164), but only one (designated 14.1) appears to be functional, possessing the three-exon structure of λ (165–167). Interestingly, a sequence upstream of Jλ1 is homologous to exon 1 of 14.1/λ5, suggesting that 14.1 and Jλ-Cλ1 may have had a common ancestral gene that could be expressed in either of two ways: (a) by rearranging its J-like exon 2 with a V region gene, like modern λ genes; or (b) without rearrangement using exon 1, with the encoded protein assembling with a non-covalently linked VpreB-like subunit. The human VpreB homolog lies within the Vλ cluster (168), in contrast to murine VpreB which lies close upstream of λ5.

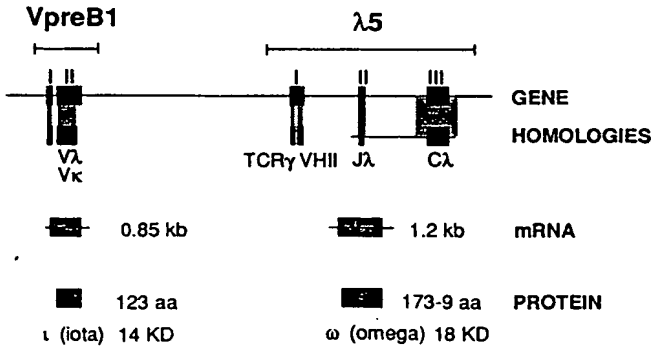


FIG. 10. λ-related genes that encode a surrogate L chain. The top line of the diagram portrays the exons of the VpreB1 gene and the λ5 gene of the mouse, which have been physically linked as shown. The second line shows sequence similarity relationships with other known Ig or TCR sequences. The expressed mRNAs and proteins that have been detected from these two genes are shown below.

V GENE ASSEMBLY RECOMBINATION

The mechanism by which germline variable-region constituent (VL and JL, or VH, D, and JH) assemble in the DNA to form complete active V gene has been pursued ever since Ig gene recombination was first discovered. In this section we address (a) the topology of the recombinations from a "macro" viewpoint, (b) the components of the recombinase machinery (a "micro" view), and (c) the regulation of that machinery in B-cell development.

Topology of V Assembly Recombination

Deletion Versus Inversion

The earliest model for Vκ-Jκ rearrangement assumed that V segments and J segments were all oriented in the same direction of transcription and that the DNA between the recombining V and J segments was simply excised and lost from the cell (Fig. 11A). However, Southern blotting of a panel of myelomas and normal κ-bearing lymphocytes showed that some cells had retained the DNA just upstream from Jκ1, a region that should have been absent from all chromosomes that underwent deletional recombination (169). Although several complex models were proposed to explain such results, the presently accepted explanation is simple: some V genes are oriented in the opposite direction from the Jκ-Cκ region. This topology would allow the VJ recombination to occur by an inversion of the DNA between the recombining V and J segment (Fig. 11B), leaving the DNA upstream from Jκ1 retained on the chromosome. The same recombinase machinery can presumably rearrange the germline elements by either inversion or deletion—depending on the relative orientations of the sequences—because

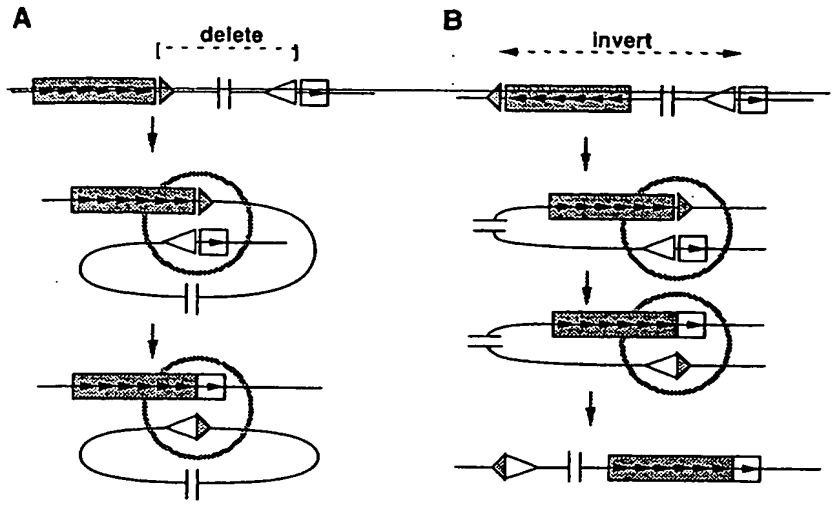


FIG. 11. The same micro mechanism of recombination can join Vκ and Jκ by deletion or inversion, depending on the relative orientation of the two precursors in germline DNA. **A:** When the V coding sequence (*shaded rectangle*) and the J coding sequence (*white rectangle*) are oriented in the same 5′→3′ direction in germline DNA (*internal arrowheads*), the recombination yields a VJ coding joint plus a DNA circle containing the signal joint (*apposed triangles*). **B:** If V is oriented in the opposite direction in germline DNA, then an identical recombination reaction at the micro level (*inside shaded circle*) leaves the signal joint linked to the recombined VJ coding joint.

this enzymatic machinery "sees" only the DNA in the immediate vicinity of the recombination site (circled in Fig. 11) and is insensitive to the topology of the DNA strands far from this site. One implication of this model (Fig. 11B) is that cells that have undergone an inversional Vκ-Jκ recombination should retain on the chromosome a recombination joint with two sets of signal sequences—the RSS from downstream of the Vκ and the RSS from upstream of the Jκ segment—joined together. Indeed such signal joints (also known as flank products and reciprocal joints) have been detected in several cell lines (170–172). In contrast to the flexibility observed in the position of the recombination breakpoint in the VJ segment (the coding joint), the sequences of signal joints usually show the J-derived 7-mer joined directly to the V-derived 7-mer, without even a single intervening nucleotide between them. Surprisingly, the signal joints retained on the expressed chromosome have almost always been derived from Jκ1. As additional evidence that inversion can occur in Vκ-Jκ recombinations, several laboratories (173,174) have reported that engineered gene constructs carrying Vκ and Jκ recombination signals in opposite orientation can undergo recombination by inversion when transfected into a B-lymphoid cell line.

The idea that some germline Vκ genes are oriented opposite to the Jκ-Cκ locus has been directly verified for the most J-proximal human Vκ gene segments and for one of the two large duplications in the human Vκ locus, as discussed in a later section. For the mouse Vκ locus, less is known about V orientation; but the observation of signal joints that are retained on the chromosome in murine cells can most easily be explained by inverted murine Vκ genes (175). On the other hand, in the H chain or lambda loci there is no evidence for inverted V genes or retained signal joints, so these loci probably recombine only by deletion.

When the recombination occurs by deletion, the model of Fig. 11A suggests that a signal joint is formed on a circular DNA molecule; such a DNA circle would not be attached to the main chromosome and, lacking an origin of replication and a centromere, would be expected to be lost from the progeny of the cell in which the recombination took place. Nevertheless, by isolating circular DNA from cells that are undergoing Vκ-Jκ rearrangement, it has been possible to detect the predicted molecules bearing signal joints (176), supporting the model.

Secondary Recombinations

A final issue for consideration of V assembly topology at the macro-level concerns secondary V gene recombinations. As discussed in an earlier section, the flexibility of VJ or VDJ joining causes nonproductive out-of-frame recombination with high frequency. A B-lymphocyte that rearranged its κ genes nonproductively on both parental chromosomes might be thought to have no further avenue for making a functional L chain; but the availability of upstream V genes and downstream J segments could allow additional recombinations to occur, as shown in Fig. 12A. More complex events are possible as a consequence of the inverted orientation of some Vκ genes. The occurrence of such secondary recombinations has in fact been reported for κ genes (172) and would be implied by the recovery of chromosomal signal joints that are not reciprocal to coding joints in the same cells. The preponderance of Jκ1-derived nonreciprocal flank products observed in myelomas may result from initial nonproductive recombinations between this J segment and inverted V genes, followed by successive recombinations involving more downstream J segments; by the time a productive rearrangement occurs, many myelomas will carry signal joint relics of earlier recombinations involving Jκ1 (177). In addition to lymphocytes with nonproductive VκJκ junctions on both chromosomes, cells that have assembled a productive VκJκ joint may undergo secondary recombination if the resulting VH-VL pair recognizes an autoantigen; this type of secondary recombination, known as receptor editing, is considered in more detail later in this chapter.

For H-chain genes the possibility of secondary recombination might seem to be ruled out by the fact that a VDJ rearrangement must eliminate all the 12-bp spaced signal elements from the VH locus because these elements are deleted on both sides of the D region that is retained in the recombined VDJ unit, and from all the germline D segments eliminated by the VD and DJ recombination events (Fig. 12B). Secondary DJ rearrangements should be possible before VHD recombination removes unrearranged upstream DH segments (Fig. 12B) and indeed this has been shown to occur (178). Of greater functional interest has been the demonstration (179) that upstream germline VH genes can recombine with an established VDJ unit, displacing the originally assembled V gene. This type of

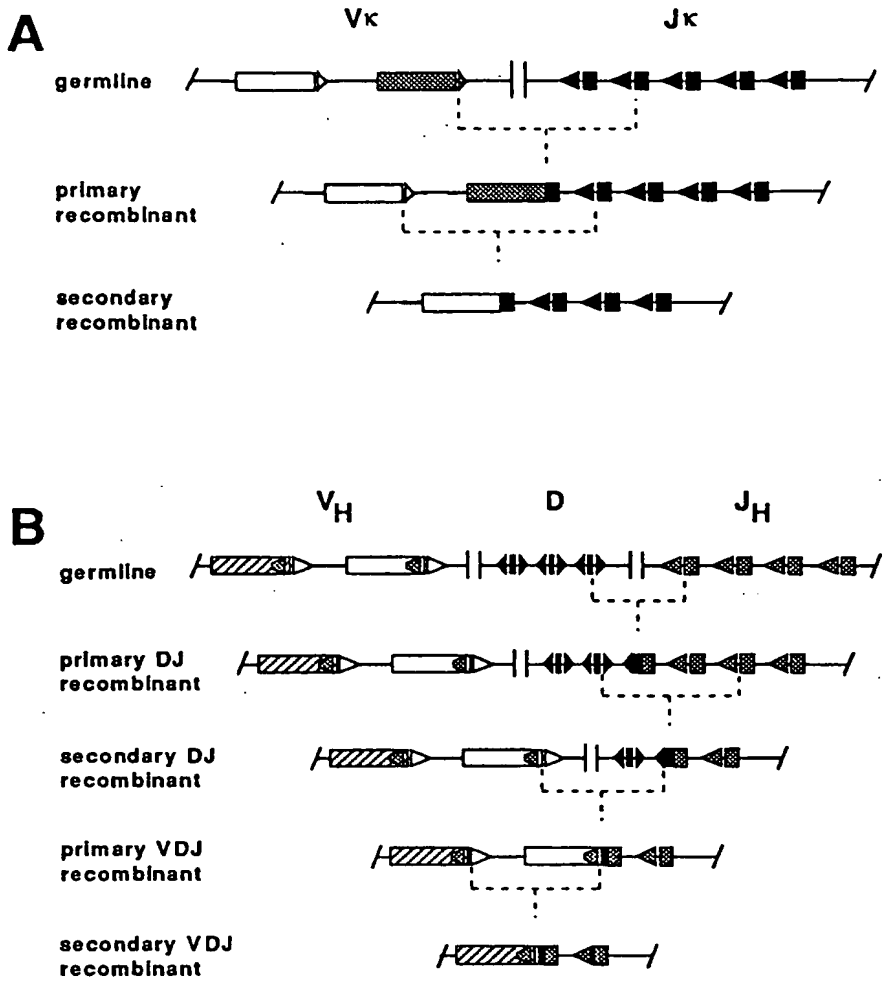


FIG. 12. Secondary recombinations. **A:** In the κ L-chain system, a primary recombination can be followed by recombination between an upstream V and a downstream J. **B:** Analogous secondary recombinations can occur in the H-chain system between upstream D and downstream J segments. After VDJ recombination eliminates all short spacer signal elements from the chromosome, secondary recombination can still occur between VH (long spacer signal) and an internal heptamer within the VH coding sequence of the VDJ unit.

recombination is apparently mediated by a sequence that closely matches the consensus signal 7-mer and that appears near the 3′ end of the coding region in about 70% of VH genes (Fig. 12B) (179,180). The internal 7-mer is not found in most L-chain genes. After VH replacement, the few nucleotides remaining from the originally assembled VH could potentially contribute to diversity; such nucleotides would be difficult to distinguish from N-region nucleotides. Secondary recombination thus represents an escape mechanism for cells with nonproductive rearrangements on both H-chain chromosomes, or, as alluded to above, for cells whose antibody encodes an autoantigen (181); it is not known, however, how frequently such escapes occur in these circumstances, as opposed to the alternative path of cell death. The fact that the isolated 7-mer is apparently able to function in VH replacement recombinations without an associated 9-mer again suggests that the 7-mer is the more critical recombination signal, although it has been suggested that an additional consensus sequence upstream of the internal 7-mer (181) may contribute to VH replacement recombination.

A Micro View of the Mechanism of V Assembly Recombinase Machinery

As mentioned earlier, the same recombinase machinery is believed to mediate V gene assembly recombinations of all four types in the Ig gene systems (κ, λ, and VH-D and D-JH) as well as similar events in the four TCR gene loci. This belief is based on the observation that all these systems share the same 7-mer/9-mer RSS and follow the same 12/23-spacer rule of recombination. Furthermore, gene constructs designed to test in vitro recombination of TCR gene segments were found to be accurately recombined when transfected into B cells. The severe combined immunodeficient (SCID) mouse strain was found to have a deficiency in recombination of both Ig and TCR genes, suggesting that both systems could be affected by a single gene defect (182). Finally, the two recombination activating genes, RAG-1 and RAG-2, have been found to be key mediators of both Ig and TCR gene recombination. The assumption that the same recombinase machinery operates on these two gene families has allowed investigators to pool knowledge concerning the mechanism of the recombination gained from studies of both systems. On the other hand the assumption of a common recombinase raises the question of how B cells preferentially rearrange Ig genes (and T cells TCR genes) when both gene systems are available to be rearranged by the common recombinase in both cell lineages; this issue will be addressed later in this section. The mechanism of V gene assembly has been investigated by several different strategies: sequence analysis of normal substrates and products (germline and recombined DNA), the use of plasmid substrate constructs capable of recombination after transfection

into lymphoid cells (in order to assess the effects of alterations in the substrate sequences), the study of presumed intermediates in the enzymatic reaction, purification of proteins that bind to RSS motifs or that perform enzymatic functions hypothesized to occur during the recombination, studies of mutations that affect the efficiency or fidelity of VDJ recombination, and, ultimately, cell-free in vitro studies using cell extracts or putative components of the recombination machinery.

Recombination Model

A model for the detailed mechanism of the recombination event must account for the observed features of the recombination products—the coding and signal joints—and of their germline precursors. The features in the germline precursors that appear necessary and sufficient for recombination are the 7-mer and 9-mer RSS with appropriate spacing (12 and 23 bp); model substrates containing these elements are competent to undergo recombination even in the absence of normal V, D, or J coding regions, although the efficiency of recombination can be influenced by features of the sequences replacing the coding regions. As for the products, the features of the signal joints are relatively simple: the 7-mers are joined "back-to-back," with only rare additions or deletions. The features of the coding joints are more complex, due to the flexibility of junctions as discussed earlier:

1. A variable number of bases are deleted from the ends of the coding regions (in comparison with the "complete" sequence in the germline precursor).
2. Nongermline nucleotides (N regions) unrelated to the germline precursor sequences are added in some coding joints; these are generally rich in G and C nucleotides.
3. Less frequently, extra bases are added that can be interpreted as P nucleotides; these are nucleotides that are joined to the end of an undeleted coding sequence and that form a palindrome (P) with that sequence end (183,184). P nucleotides are generally only 1 or 2 bp, but they can be longer, especially in mice with the SCID defect (185).

The recombination model first proposed by Alt and Baltimore (186) accommodates many of these observations and, with some recent modifications, can serve as a framework for consideration of the recombination mechanism (Fig. 13). The recombination is thought to begin with binding of components of the enzymatic recombinase machinery to the 7-mer–9-mer RSS adjacent to the two segments to be recombined. Both DNA segments are then cleaved at the border of the two 7-mers (a reaction now known to be catalyzed by the RAG genes). The two 7-mer ends are joined together without modification, but the ends that will form the coding joint (which are now known to exist transiently in the form of a "hairpin" loop) are digested to varying extents by an exonuclease activity. Variable numbers of nucleotides may be added to the 3′ ends through the action of terminal deoxynucleotide transferase (TdT). Then the 5′ ends are filled in by a DNA polymerase and the resulting flush ends are ligated together, completing the recombination event.

In Vitro Experiments to Investigate Substrates and Products

Investigations of the recombination mechanism have been advanced by the development of methods for following these events in vitro. Some experiments have exploited the ability of the Abelson murine leukemia virus (AMuLV) to selectively transform pre-B cells without abolishing the active V gene assembly characteristic of this stage of lymphoid development. Several AMuLV lines have been cloned and then repeatedly subcloned in order to follow the progression of recombination events (187–189). Particularly valuable information has been gained by transfecting AMuLV lines, as well as other lymphoid and nonlymphoid cells, with artificial gene constructs capable of undergoing V(D)J recombination. Several such constructs have been designed with selectable markers whose expression depends on a recombination event. For example, Lewis et al. (190) used a retroviral construct in which a drug-resistance gene was placed between Jκ and Vκ sequences such that the gene could be expressed only after inversional VJ recombination. In another strategy, Lieber et al. (191) transfected various cell lines with a plasmid containing an ampicillin resistance gene (Amp^r) plus a chloramphenicol resistance gene (Cam^r) whose expression was blocked by a stop codon flanked by two V(D)J recombination signal sequences. In B cells, recombination between the two signal sequences deletes the stop codon, allowing expression of the Cam^r gene. When extrachromosomal circles are recovered from the cells and transfected into bacteria, the extent of recombination can be determined by the ratio of transformed Amp^r bacteria that are Cam^r. Depending on the orientation of the signal sequences in the starting construct, the recombination products represent coding or signal joints. These joints can be recovered efficiently from the Cam^r colonies for analysis; the sequences of these joints have all the characteristics of natural recombination products. One interesting outcome from such experiments was the discovery that pre-B and pre-T cells from SCID mice were capable of recombination to form signal joints but were markedly defective in their ability to join coding ends to form coding joints (192).

Indeed, from recombined engineered substrates, certain nonstandard joints also have been recovered, which, although not contributing to physiologic V gene assembly, may reflect features of the recombination mechanism (193). These nonstandard joints can be understood by appreciating that there are three topologies in which DNA that has been cut twice—generating four ends—can be rejoined. If the four ends are coding(V), signal(V), signal(J), and coding(J), the three possibilities can be defined by considering the three different ends that may join to the coding(V) end (assuming that the remaining two ends must join to each other). The possibilities are as follows:

1. coding(V)–coding(J) plus signal(V)–signal(J). This is the standard reaction product in which the coding(V)–coding(J) product encodes the assembled VJ gene and the signal(V)–signal(J) represents the signal joint.
2. coding(V)–signal(V) plus signal(J)–coding(J). These products (open and shut joints) look like the starting DNAs but can be distinguished from them if nucleotides have been added or deleted at the junctions so that they no longer hybridize to oligonucleotide probes specific for the coding/signal junction.
3. coding(V)–signal(J) plus signal(V)–coding(J). These are hybrid joints, in which the signal ends have switched places.

The fact that all these recombinations can occur readily in the transfected construct DNA—which contains little Ig gene sequence beyond the DNA immediately flanking the V and J—suggests that neither a specific chromosomal location nor extensive flanking sequences are necessary for the recombination.

The critical characteristics of the signal sequences have been explored using transfected constructs carrying various mutations (194,195). These experiments have verified the importance of the 7-

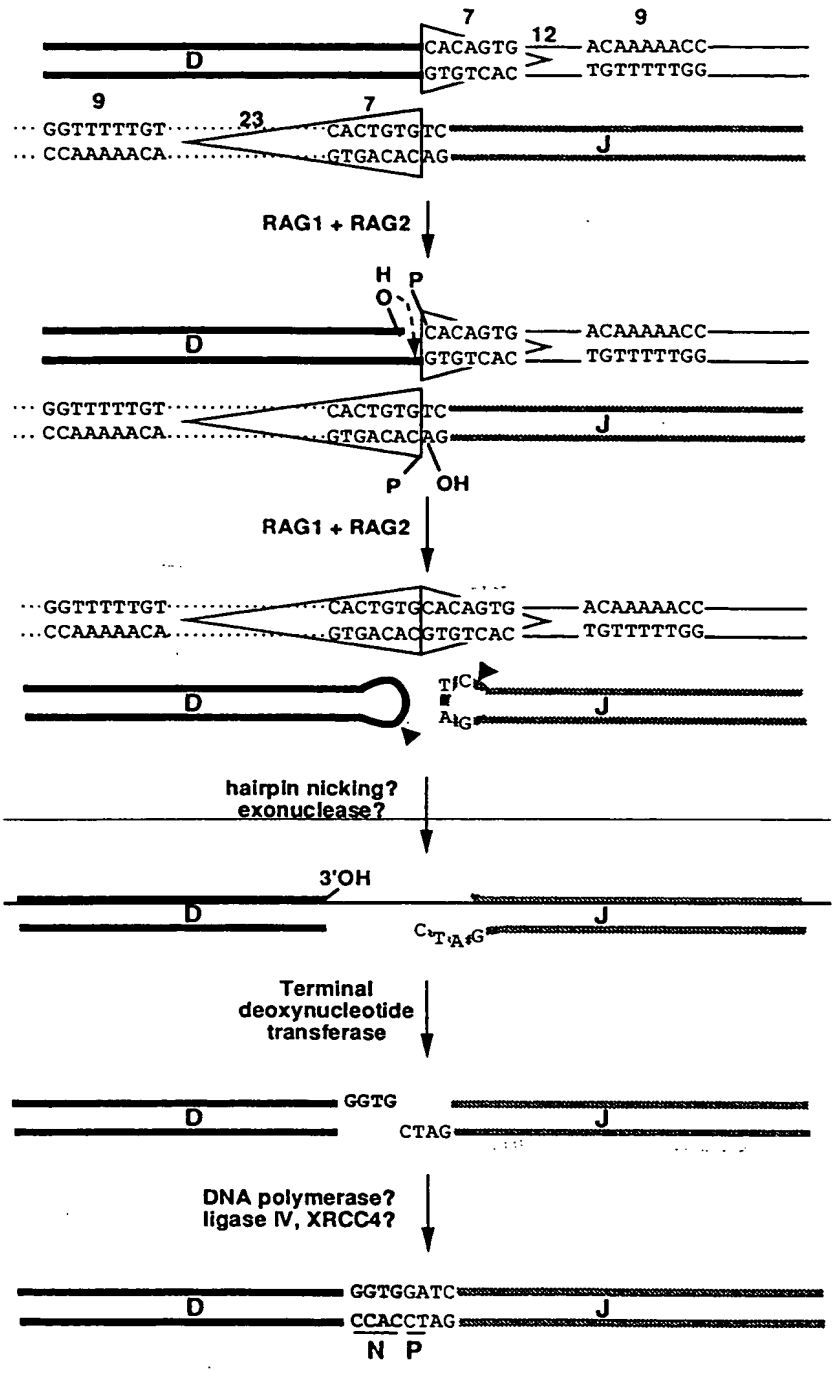


FIG. 13. Model for V assembly recombinations. All V assembly recombination reactions (in Ig and TCR genes) may proceed by a common mechanism, illustrated here by D-J recombination. The RSS 7-mers are depicted in triangles, which is the conventionally used RSS graphic. Hairpin loops are created on coding ends dependent on the action of RAG1 and RAG2.

This reaction also generates two signal ends, which are ligated together. In the example shown here, after the opening of the hairpin loops on the coding ends, the D coding sequence is nibbled by exonuclease, whereas the J coding sequence is spared and instead shows P nucleotide generation due to asymmetric hairpin cleavage. N-region addition is pictured in this example as occurring only on the D-region end, but in reality, exonuclease digestion and N nucleotide addition can occur on either (or both) coding ends. The steps in the proposed mechanism are discussed in the text.

mer/9-mer sequences and have shown that the spacer sequences are not critical as long as the spacer length (12 or 23 bp) is maintained. The 7-mer appears especially important, with the 3 bp closest to the coding sequence being most critical for signaling recombination. Recombination activity could be detected in a variety of pro-B and pre-B cells as well as pre-T lines, but was virtually absent in mature B and T cells and in cells of nonlymphoid lineages (191).

P Regions and Hairpin Intermediates

To study intermediates in V(D)J recombination, Roth et al. (196,197) designed a Southern blot strategy to detect double-strand breaks in the TCR δ locus between D and J, and applied this strategy to DNA from newborn murine thymus cells, which actively rearrange the TCR δ locus. Compared with DNA from adult liver, additional bands were found in the newborn thymus DNA, representing DNA fragments extending from a D-derived 7-mer on one end to a J-derived 7-mer on the other end. In order to characterize the sequence of the signal ends in detail, several laboratories have used ligation-mediated PCR (LM-PCR). This technique involves ligating the blunt genomic signal ends with a double-stranded oligonucleotide, followed by amplification extending from a primer sequence near the genomic signal end to the added oligonucleotide sequence; amplified products can then be cloned and their sequence determined. LM-PCR analyses of both TCR and Ig genes have defined the signal ends as blunt-ended cuts, usually exactly at the 7-mer border, leaving 5′ phosphate and 3′ hydroxyl groups (198,199).

In their original Southern blot assays, Roth et al. detected the signal ends from these cuts at levels representing about 2% of the thymus DNA; but the coding ends could not be visualized at all, perhaps because of rapid processing of these ends into coding joints. Based on the known defect of SCID lymphocytes in forming coding joints, Roth et al. (196) reasoned that SCID thymocytes might accumulate the cut coding ends that could not be visualized in normal thymocytes. Indeed, coding ends were detected in the SCID thymocyte DNA and, moreover, were found to have several properties suggestive of a hairpin-like structure. First, the coding ends in SCID thymocyte DNA were resistant to exonuclease treatment. Furthermore, restriction fragments bearing these in vivo–generated ends on one side were found to move on a denaturing electrophoresis gel as if they were twice as long as predicted from the size of the double-stranded fragment before denaturation. Finally, LM-PCR experiments failed to detect the coding ends unless they were pretreated with a single strand–specific endonuclease, consistent with the impossibility of ligation to a hairpin unless it was first opened. The sequences of LM-PCR products obtained after endonuclease treatment suggested that the hairpins contained the entire sequence of the coding element, without loss or gain of a single nucleotide (200).

One interpretation of these hairpin ends in the SCID DNA is that they represent normal V(D)J recombination intermediates that—in wild-type cells—are opened at variable positions within the hairpin loop by an endonuclease activity that is dependent on the normal allele of the SCID gene. P nucleotides could then result from opening the loop at an asymmetric position (Fig. 13); this model would explain the absence of P nucleotides from coding ends that have been "nibbled" after opening of the hairpin. The unusually long P nucleotide segments observed in the rare coding joints assembled in SCID mice might then be interpreted as resulting from resolution of hairpins by nonspecific nicking enzymes that, unlike the exonuclease activity dependent on the normal allele of the SCID gene, do not focus on the hairpin loops but nick in variable positions in the double-stranded hairpin "stem" (196).

In support of the notion that the hairpin coding ends are intermediates in normal VJ recombination, Ramsden and Gellert (201), using LM-PCR, were able to detect such ends in a normal (non-SCID) B-lymphoid line engineered to sustain a high level of κ gene recombination. In this line, broken Jκ1 signal ends were detectable in amounts corresponding to 30% to 40% of the κ loci present. Coding ends were observed at 10- to 100-fold lower abundance, with both hairpin and open ends. The observed kinetics were consistent with simultaneous production of signal ends and hairpin coding ends, with rapid processing of hairpins to open coding ends and then to coding joints, but slower ligation of signal ends. This model is further supported by the observation that linear DNA molecules with hairpins at both ends can, after transfection into B-cell lines, be recovered as recircularized molecules, with the frequent creation of P insertions (202). Interestingly, SCID B cells perform about as well as normal B cells in this assay, suggesting that the protein missing in SCID cells is not the hairpin nicking enzyme itself, but rather an activity that makes natural endogenous hairpin coding ends available to the enzyme; these endogenous ends may require such an activity because of their association with recombinase or other chromosomal proteins, whereas transfected hairpins free of attached proteins may be accessible independent of the protein missing in SCID. The molecular basis of the SCID defect is considered in more detail below.

In Vitro Cell-Free VDJ Recombinase Activities

In attempts to discover components of the VDJ recombination machinery, several groups have studied endonuclease activities from lymphoid sources that cleave DNA selectively near the 7-mer element (203–205); but these experiments have not led to any breakthroughs, and it presently seems unlikely that any of these activities represent components of the recombinase.

RAG Genes

A rather unlikely experiment has led to a major breakthrough: the identification of two genes whose products are apparently critical for V(D)J recombination in B- and T-lineages. Schatz and Baltimore (206) stably transfected fibroblasts with a construct containing a selectable marker whose expression was dependent on VDJ recombination; as expected, no measurable recombination occurred in this nonlymphoid cell. However, when either human or murine genomic DNA was transfected into these fibroblasts, a small fraction of recipient cells stably expressed recombinase activity, activating the selectable marker. This suggested that a single transfected genomic DNA fragment was able to confer recombinase activity in a fibroblast. After successive rounds of transfection and selection for recombinase activity, the critical genomic fragment was identified. This fragment turned out to contain two genes, designated RAG-1 and RAG-2. These genes are not homologous to each other, and neither is strikingly similar to any other known genes, although a weak relationship between RAG-1 and a topoisomerase has been suggested. Both genes are required for activity; therefore, these genes would not have been discovered by this transfection technique if they had been situated too far apart in the genome for both to be transferred on a single DNA fragment. The genes are notable for having no introns in most species (certain fish are exceptions) and for their close association and oppo-

site transcriptional orientation in all species examined. The close proximity of these two genes related by function but not by sequence has led to the speculation that they might have arisen from a more primitive viral or fungal recombination system (207), as discussed later in this chapter.

A crucial role for the RAG genes in V assembly recombination is supported by the conservation of these genes in a variety of Ig-producing vertebrate species, from humans to sharks (208), whereas RAG homologs have not been identified in any species that does not demonstrate Ig V gene assembly recombination. RAG-1 and RAG-2 are expressed together in pre-B and pre-T cells, specifically at the stages expressing V(D)J recombinase activity. Moreover, mouse strains in which either gene has been eliminated by homologous recombination (gene knockouts) have no mature B or T cells as the apparent result of abrogation of V(D)J recombination (209,210). Recently a subset of human patients with a SCID syndrome and no B-lymphocytes were found to have mutations in RAG genes (211).

Attempts to demonstrate activities of the RAG proteins on recombination substrates in vitro were hampered by poor solubility of the proteins, but functional analyses of mutated RAG genes—using RAG expression vectors cotransfected with recombination substrate plasmids into fibroblasts—showed that surprisingly large segments of both proteins can be deleted without eliminating recombinase activity (212); and some of the deleted proteins were soluble and could be handled relatively easily as fusion proteins. This work allowed the demonstration that in a cell-free in vitro system the two RAG proteins together can perform cleavage of substrate DNAs as well as hairpin formation on the coding end (213). The reaction occurs in two steps: first a nick occurs on one strand adjacent to the heptamer—the top strand as drawn in Fig. 13—then the 3′ hydroxyl created at the nick causes transesterification by nucleophilic attack on the phosphodiester bond adjacent to the 7-mer on the bottom strand (Fig. 13), yielding a hairpin on the coding end and a new 3′ hydroxyl on the 3′ end of the bottom 7-mer strand (214). This transesterification mechanism is consistent with the observation that the formation of the new phosphodiester bond in the hairpin occurs in the absence of external energy source such as ATP, apparently using the energy inherent in the phosphodiester bond broken in the nucleophilic attack. The stereochemistry observed in the reaction suggests that no phosphodiester linkage to protein occurs as an intermediate, such as occurs in bacteriophage lambda integration; instead the direct transesterification mechanism resembles the μ transposition and retroviral integration, which can both produce hairpins under certain experimental conditions (214).

The actions of the RAG proteins were found to be dependent on divalent ions in the medium (215,216). In Mn^{2+} the RAG proteins catalyzed cleavage of substrates with a single RSS; but in Mg^{2+} cleavage required two RSSs and occurred most efficiently if the substrate conformed to the 12/23 rule regarding the spacing between 7-mer and 9-mer elements; thus, this rule may be a result of RAG protein specificity, although other proteins also seem to contribute to 12/23 specificity, perhaps by promoting an optimal molecular architecture (217). In Ca^{2+} the RAG proteins and a radiolabeled DNA substrate containing an RSS formed a stable complex that was apparent in an electrophoretic mobility shift assay (EMSA) and was stable to competition with unlabeled substrate, but the substrate was not nicked or cleaved. However, when Mg^{2+} was added to the stable complex, substrate cleavage occurred. (Interestingly, in the presence of Ca^{2+} human immunodeficiency virus (HIV) integrase and μ transposase also form similar stable complexes in which substrate DNA is bound but no cleavage occurs.) The Ca^{2+}-mediated binding of RAG proteins to substrate DNA was decreased 10-fold by the elimination of the 9-mer from the RSS, so the RAG proteins must recognize both components of the RSS. In contrast, mutations in the 7-mer that altered the nucleotides closest to the coding region (residues known to be critical for supporting cleavage) had minimal effect on binding. This is consistent with other evidence (218,219), suggesting that these nucleotides, and the adjacent nucleotides in the coding region, may contribute to a local alteration in DNA helix structure that is important for the cleavage reactions. Recently, in vitro experiments have been reported in which RAG proteins supplemented with extracts from several cell lines were able to generate signal joints (220) and coding joints (221,222). In one system, RAG-1 and RAG-2 could be detected in a stable complex containing two signal ends, an HMG (high-mobility group) protein, and perhaps other proteins as well (223). These in vitro VDJ recombination experiments should allow the elucidation of other components required for the reaction, as well as its mechanism.

The double-strand DNA breaks catalyzed by the RAG proteins could be potentially deleterious if they occurred during DNA synthesis or mitosis, but this problem appears to be prevented by tight posttranscriptional regulation of RAG-2 protein levels across the cell cycle. Although the RAG-1 protein and mRNA transcripts of both RAG genes vary little across the cell cycle, a phosphorylation-dependent degradation signal mediates destruction of the RAG-2 protein (224,225), thereby preventing double-strand DNA breaks in the H-chain JH locus from occurring during M, G2, and S (199). The phosphorylation site, a threonine at amino acid 490, falls into a region of the sequence that is highly conserved across species and contains a consensus sequence characteristic of targets of cyclin-dependent kinases. This regulatory region is dispensable for enzymatic activity. In RAG-2 knockout mice carrying a transgenic RAG-2 gene with an alanine replacing the phosphorylatable threonine, RAG-2 protein and double-stranded DNA breaks were found throughout the cell cycle, demonstrating the importance of the RAG-2 degradation signal in cell cycle control of VDJ recombination (226).

Although all the binding and enzymatic activities of the RAG proteins discussed so far have required the presence of both RAG-1 and RAG-2, recent reports have suggested that RAG-1 may bind weakly to the RSS 9-mer in the absence of RAG-2 and that this binding may be mediated by a segment of RAG-1 that bears sequence similarity to the DNA binding domain of bacterial invertases (227,228). Two circumstances have been described in which only one of the two RAG genes is expressed. RAG-1 was reported to be expressed without RAG-2 at low levels in the developing central nervous system (229) (although the RAG-1–deficient mice show no obvious central nervous system defects). Conversely, RAG-2 is expressed without RAG-1 in the chicken bursa of Fabricius (230), which contains B-lineage cells at a developmental stage when their genes have already undergone V(D)J rearrangement and are in the process of being diversified by gene conversion (as discussed later in this chapter). The significance of this finding is unclear at present because RAG-2 is apparently not essential for the gene conversion itself (231).

Apart from the obvious importance of the RAG proteins in mediating the initial steps of VDJ recombination, knowledge of these proteins and their genes has allowed two major technical advances that have opened the way to many additional experiments. One such advance is the availability of the RAG-1 and RAG-2 knockout mice. These mice have no functional B cells or T cells, and are not "leaky" like SCID mice, which develop some

functional B and T cells, especially as the animals age. The RAG knockouts can be used to study the importance of the innate immune system (i.e., responses that occur in the absence of antigen-specific lymphocytes) in particular immune responses. The knockouts can be used as recipients for various lymphocyte subsets to explore the roles of different cell types. They can be used to study the signals for B cell development by introducing transgenes with specific functionally recombined Ig genes and characterizing the phenotypes of lymphocytes that develop. Finally, they can be used in RAG complementation experiments designed to assess the phenotype (in lymphocytes) of various other gene knockouts (232). In RAG complementation, embryonic stem (ES) cells in which the gene of interest has been knocked out by homologous recombination are injected into homozygous RAG knockout blastocysts (RAG-/-); this procedure yields chimeric mice in which all B and T cells derive from the engineered ES cells, which are the only source of intact RAG genes to support lymphocyte development. Such animals can be made more easily than a knockout mouse line and can be used to study the effect of gene deletion in lymphocytes independent of effects the deletion may have in other cells. In particular, for cases where the gene knockout causes embryonic lethality due to effects on nonlymphoid cells, RAG complementation allows the selective knockout in lymphocytes to be studied.

The second major technical fallout from the RAG genes is the method for investigating VDJ recombination in nonlymphoid cell lines with well-characterized mutations in genes governing DNA repair; when such lines are transfected with RAG genes, the effects of these gene mutations on VDJ recombination can be assessed.

Components of Later Steps in VDJ Recombination

Clearly the RAG genes are critical for the first steps in VDJ recombination (recognition of RSS, cleavage, and hairpin formation), but additional components are required to complete the reaction; and at least some of these components may function not only in VDJ recombination but also in ubiquitous DNA repair pathways. The first clear example of such a component to be recognized was the murine SCID mutation described above. This mutation was originally identified in a mouse strain that was immunodeficient as a result of a marked impairment in VDJ recombination of both Ig and TCR genes; SCID lymphocytes can perform the RAG-mediated reactions of cleavage and hairpin formation, and can form signal joints, but are markedly defective in coding joint formation. Subsequently it was found that the SCID mutation also blocks the enzymatic mechanism responsible for repair of double-strand DNA breaks—such as those caused by ionizing radiation—in both lymphoid and nonlymphoid cells. This suggested that after RAG-mediated DNA cleavage, lymphocytes may complete the joining reactions using enzymes that function ubiquitously in DNA repair. To test this hypothesis, Taccioli et al. screened panels of Chinese hamster ovary (CHO) cell lines carrying well-characterized defects in DNA repair to see if these lines were impaired in performing VDJ recombination after transfection with the RAG genes (233). Such cells had previously been classified into x-ray cross complementation (XRCC) groups by investigating the outcome when two mutant cell lines are combined to make a somatic hybrid. If such a hybrid shows no DNA repair defect, this implies that the two original cell lines carry different mutations such that the hybrid ends up with a normal copy of each gene (the mutant cell lines cross-complemented one another). Conversely, by definition, cells in the same cross-complementation group are unable to complement each other. Of eight XRCC groups of ionizing radiation-sensitive rodent cell lines, three were known to be defective in repair of double-strand breaks in DNA (XRCC groups 4, 5, and 7), and all three of these groups were found to be impaired in VDJ recombination after RAG transfection.

The genes mutated in XRCC 5 and XRCC 7 turned out to encode two components of a three-polypeptide complex known as the Ku complex. Originally characterized as the autoantigen recognized by a patient antiserum, Ku is composed of an approximately 70-kDa protein (Ku70) and an approximately 86-kDa protein (Ku86, often called Ku80) which, when heterodimerized, can bind to DNA (234). Then the DNA-Ku heterodimer complex can recruit the third component: an approximately 450-kDa protein with a protein kinase activity that is dependent on binding to DNA (234,234a). Some evidence suggests that DNA-PK can bind DNA even in the absence of Ku (234b). The conservation of Ku genes in drosophila and yeast suggest that the complex evolved long before VDJ rearrangement. Extensive evidence indicates that the gene defective in XRCC 5 encodes Ku80 (235,236), whereas the gene defective in XRCC 7 (which corresponds to the gene mutated in murine SCID) encodes the 450-kDa DNA-PK (237). Although the murine SCID mutation impairs primarily coding joints, this difference is probably a result of residual DNA-PK protein in the SCID cells because a more complete equine DNA-PK mutation impairs both coding and signal joints (238). Ku80 mutant cell lines are also defective in both signal and coding joint formation, as are mice with a knockout of the Ku80 gene (200). Ku70 mutants have not been detected in panels of existing XRCC mutants, but recent evidence indicates that cells with homozygous disruption of Ku70 are also defective in VDJ recombination induced by RAG gene transfection (239).

The Ku complex had previously been studied as an activity with an unusual DNA binding specificity: rather than recognizing particular nucleotide sequences in DNA, it recognizes particular topologic features of DNA, including double-stranded DNA ends (such as might be generated by double-strand breaks caused by x-rays or by recombinases). Once bound to an end, it can translocate down the length of the DNA (240). The Ku heterodimer also has been reported to have an ATP-dependent helicase (DNA-unwinding) activity (241). Several models can be envisioned for the role of this complex in VDJ recombination: the complex may bind to the hairpin coding ends and regulate hairpin opening and DNA degradation by exonucleases; it may participate in destabilizing the DNA double-helix through helicase action; and it may influence recombination by phosphorylating other proteins via the protein kinase activity of DNA-PK. Additional investigation will be necessary to clarify which (if any) of these roles is important for VDJ recombination.

The gene mutated in XRCC group 4 also has been cloned (242). It encodes a ubiquitously expressed protein of about 38 kDa predicted size that is not homologous to any known protein. The protein has recently been found to bind to and activate DNA ligase IV, suggesting that this enzyme is probably important for ligating signal and coding joints in VDJ recombination (243,244). In addition, the XRCC4 protein product interacts with DNA-PK and is phosphorylated by this kinase (244a).

Other Proteins that May Participate in V(D)J Recombination

The RAG proteins, the Ku-DNA-PK complex, and the XRCC 4 protein have all been shown to play a role in VDJ recombination because mutations that compromise these proteins impair the recombination. Various other entities have been proposed as participants in VDJ recombination; some of these are described below, although it is not clear that any of them participate in VDJ recombination.

Several laboratories have sought components of the recombinase machinery by searching for proteins that (a) are present in nuclear extracts from cells early in the B-lineage in which VDJ recombination is occurring and (b) bind in vitro to the 7-mer or 9-mer signal sequences in a sequence-specific manner. Although this would seem a reasonable approach, it is clear that binding to RSS sequences does not necessarily imply a physiologic role in the recombinase reaction. The RAG proteins, which clearly bind to RSS sequences, seem both necessary and sufficient to initiate VDJ recombination, so it is possible that no further RSS-specific components are necessary for the recombinase. An instructive cautionary example is provided by the case of the RSS-binding protein for Jκ (RBP-Jκ). This protein was purified from a murine pre-B cell line on the basis of sequence-specific binding in vitro to a Jκ 7-mer, was found in lymphoid cell lines but not in nonlymphoid lines, and bears sequence homology to bacterial integrases, which reinforced its possible role in a DNA recombination; but more recent studies suggest a function for RBP-Jκ that is unrelated to VDJ recombination (245).

With such caveats in mind, several other candidate participants in VDJ recombination can be mentioned. One protein capable of binding to a probe containing the 7-mer signal element was detected by EMSA in several pre-B cell lines, but not in myeloma, mature T-cell; monocyte, or fibroblast cell lines, consistent with a role for the protein in recombination-competent cells (246). Another RSS-binding protein has been identified by a technique known as Southwestern analysis: a protein extract is subjected to SDS-polyacrylamide gel electrophoresis, blotting onto nitrocellulose, and probing with a radioactive oligonucleotide including the RSS. By this method a 115-kDa protein was detected in extracts of immature B- and T-cell lines, but not in a myeloma line; the binding of this protein seemed markedly reduced when mutations were introduced into either the 7-mer or the 9-mer in the probe (247). Another protein, designated T160, was obtained as a cDNA clone from a protein expression library that was screened with an RSS probe having a 12-bp spacer (248). The T160 gene product, expressed as a fusion protein, was able to bind to the original screening probe but not to a probe with a mutated 7-mer or to a probe with a 23-bp spacer. Another protein has been identified as binding to the 9-mer RSS element but not to several mutated versions of the 9-mer (249). Designated NBP (nonamer binding protein), this 63-kDa protein was purified approximately 20,000-fold from calf thymus. A possibly related protein designated VDJP was identified from a lymphoid cDNA expression library by screening with a Jκ RSS probe (250). The resulting full-length cDNA represents a lymphoid-specific alternative splice form of the ubiquitous replication factor C (RF-C) mRNA; both sequences contain a region homologous to bacterial ligases. In vitro–expressed VDJP protein catalyzes a DNA joining reaction dependent on an RSS sequence in the DNA fragments that are joined (251). However, the substrates and products of this joining reaction differ in many critical respects from Ig genes, so the relevance of this protein to VDJ recombination is uncertain. A protein designated recognition component (Rc) is encoded by a cDNA that was isolated from a cDNA expression library from mouse thymocyte RNA using a radiolabeled RSS as a probe (252). The in vitro–expressed protein binds DNA as multimers, suggesting a possible role in bringing together elements to be joined by VDJ recombinase machinery (253). A 30-kDa protein recognizing both 7-mer and 9-mer of the RSS was detected in immature thymocytes enriched for pre-T cells undergoing VDJ rearrangement (254). A possibly related cDNA clone with homology to DNA helicases, designated lymphoid-specific helicase (lsh), was cloned from fetal thymus (255).

Exonuclease

Many recombined V regions are found to be missing variable numbers of nucleotides at the recombination junctions compared with the coding sequences present in their respective germline V, D, or J precursors. This observation has been proposed to result from exonuclease-induced nibbling of the cut DNA ends during the time between cleavage near the 7-mer RSS and rejoining of the cut DNA ends. Although the responsible exonuclease has been sought (256), and several exonucleases are known to exist in mammalian cells, the specific enzyme that nibbles the ends of V, D, and J segments has not been definitively identified.

N Regions and TdT

Terminal deoxynucleotide transferase, the proposed source of N-region additions, is an enzyme found in thymus and bone marrow, and also is a distinguishing characteristic of lymphoid versus myeloid leukemias. It catalyzes the addition of nucleotides onto the 3′ end of DNA strands. Although no template specificity determines the nucleotides added, the enzyme adds dG residues preferentially. This fact is consistent with a role for this enzyme in the origin of N regions found at the V-D and D-J junctions because these N nucleotides tend to be G-rich at the 3′ ends of both the upstream coding strand and the downstream noncoding strand. Both N-region addition and TdT are characteristically absent from fetal lymphocytes (257). N-region addition is common in H-chain genes but rare in murine L-chain genes, although perhaps less rare in humans (258).

The proposal that N regions result from the action of TdT has received considerable support. Lymphocytes with engineered defects in their TdT genes produced rearranged Ig V regions with almost no N additions (259,260). Conversely, when TdT expression was engineered in cells undergoing κ or λ L-chain rearrangement, the normally low level of N-region insertion in these recombinations was dramatically increased (261–263). This result suggests that the low frequency of N-region sequences in normal κ or λ recombinations is not due to the inability of these coding sequences to accept N-region nucleotides. Instead, the preferential occurrence of N regions in H-chain versus L-chain genes (in mice, at least) reflects TdT levels that are higher in early B-lineage cells undergoing H-chain rearrangement than in the later stage of L-chain recombination; indeed, mice with an engineered mutation that allows premature Vκ-Jκ joining in pro-B cells show an increased frequency of N-region nucleotides in their recombined Vκ genes (264). In normal mice the expression of a μ H chain may downregulate TdT expression (265), contributing to the reduced level during the stage of L-chain recombination.

N regions are also observed in TCR genes, in which they may be particularly significant as a source of sequence diversity in view of the lower germline V diversity and absence of somatic mutation in the TCR gene systems. Although N regions clearly enhance the diversity of Ig and TCR V regions, mice lacking TdT show no significant deficiencies in immune responses (266). The normal phenotype of such mice (apart from the absence of N regions) and the lymphoid-specific expression of TdT in normal mice both support the view that the only function of this enzyme is to diversify V-region genes. In TdT mutant mice, as well as in normal fetal lymphocytes low in TdT activity, absence of N-region addition is associated with an increase in the frequency of recombination junctions in which short stretches of nucleotides could have derived from

either germline element because of an overlap of identical sequences at the coding ends. These junctions suggest a recombination intermediate in which the complementary single-stranded regions from the two coding ends hybridize to each other, much as "sticky ends" generated by restriction endonucleases can facilitate ligation of DNA fragments. Such homology-mediated recombination may restrict the diversity of neonatal antibodies. The resulting antibodies possibly are enriched in specificities for commonly encountered pathogens, or have broadened specificity, as has been reported for TCRs lacking N regions (267).

Regulation of V(D)J Recombination

The recombination events that occur among Ig genes must be among the most important events that mark the development of a B-cell clone. Regions of the genome are irrevocably deleted, and commitments are made as to which L-chain isotype and which VL-VH pair will be expressed in subsequent progeny cells. It would be expected that such significant events would be well regulated. Indeed, the observation that each B-cell line generally expresses only one L-chain isotype (isotype exclusion) and uses only one of the two homologous chromosomal loci for H- and L-chain genes (allelic exclusion) implies some form of regulation. Isotype and allelic exclusion ensure that each lymphocyte expresses a single H_2L_2 combination and thus a single antigen-binding specificity, a crucial feature of the clonal selection model of the immune response. Furthermore, if the same recombinase machinery mediates the V-gene assembly reactions of all the Ig and TCR gene systems, then some mechanism must regulate which gene systems are susceptible to recombination in B-cell versus T-cell development. Current evidence suggests that VDJ recombination is controlled at two levels: regulation of the RAG protein levels and regulation of accessibility of the recombinase machinery to the germline substrates of rearrangement. Because RAG expression and locus accessibility are in turn regulated depending on the stage of B-cell development, a brief scheme of this development is presented below as background; a detailed account is provided in Chapter 6.

B-Lymphocyte Development

Figure 14 illustrates a scheme of B-lymphoid development as elucidated by the following:

1. Analysis of lymphoid malignancies or virally immortalized cells representing different stages of arrested development
2. Purification of subpopulations of normal cells from lymphoid organs by fluorescence-activated cell sorting (FACS), followed by analyses of different subsets

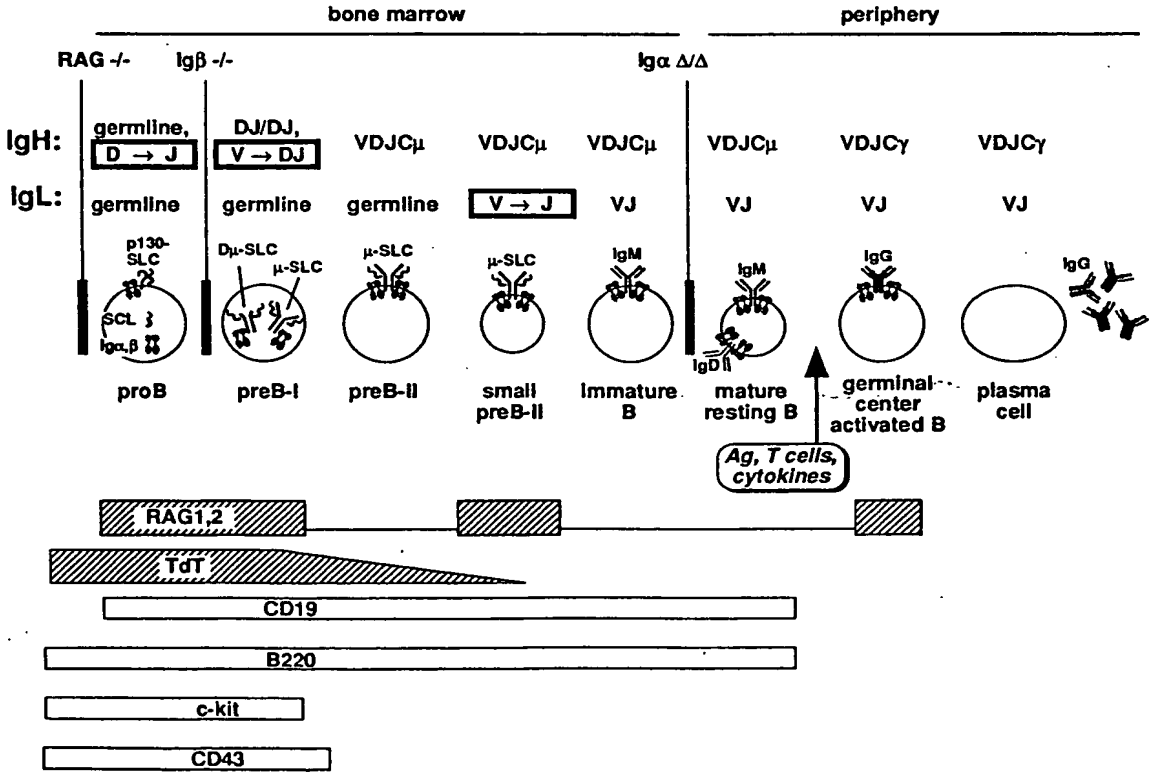


FIG. 14. Ig gene recombination in B-cell development. A simplified scheme of B-cell development is presented as a background for discussion of Ig gene recombination. The stages occurring in the bone marrow versus in the periphery (e.g., lymph nodes, spleen) are shown, along with the status of IgH and IgL genes at each stage. A graphic image depicts the Ig-related proteins displayed on the surface at each stage; at the bottom, the stage-dependent expression of RAG genes and TdT—both important in V(D)J recombination—is schematically depicted, as is the expression of several other marker proteins.

3. Phenotypic analysis of mutant mice with defects in various genes critical for progression from one developmental stage to the next
4. Culture of B-cell precursors in vitro using systems that allow developmental progression.

This figure attempts a consensus of schemes from two principal laboratories (268,269). Some known differences between human and mouse B-cell development and surface markers (270) are not reflected in this simplified summary figure.

B- and T-lymphocytes differentiate from pluripotent hematopoietic stem cells in the fetal liver and bone marrow. The primordial lymphoid progenitor has the potential to differentiate into B- or T-lymphocytes or natural killer (NK) cells. Among the earliest markers that indicate B-lineage specificity are the non-Ig components of the pre-B cell receptor (pre-BCR): Igα, Igβ and λ5 (271). CD19, which functions as a coreceptor in signal transduction, first appears in large proliferating pro-B cells, which also express several other distinguishing surface markers, including c-kit (receptor for the stem cell [growth] factor [SCF]), B220 (a B-lineage form of the phosphatase CD45), TdT, and CD43 (a sialoglycoprotein known as leukosialin). In the absence of H-chain protein, the SLC is displayed on the surface membrane in association with a complex of glycoproteins, represented by a hook shape in Fig. 14, which has sometimes been called a surrogate H chain (272). The next stage, the pre-B cell, is marked by expression of the RAG genes and H-chain rearrangement, as well as loss of c-kit and then CD43 expression. Initially DJ rearrangements occur on both chromosomal loci, producing the preB-I cell; then germline V regions join to complete a VDJ gene (preB-II). The resulting μ protein appears on the B-cell surface along with the SLCs in a pre-BCR or μ-SLC complex that also includes Igα and Igβ. The resulting large pre-B cells proliferate, with RAG gene expression downregulated. Then the cells become smaller, stop dividing, turn up RAG gene expression once more, undergo L-chain rearrangement, and express surface IgM (immature B cells). When they eventually also express surface IgD they become mature B cells and migrate into the periphery, ready to be triggered by antigen exposure.

Recombinational Accessibility and Transcription

What maintains the locus specificity of VDJ rearrangement—i.e., why is Ig gene recombination confined to B cells, with H-chain rearrangement before L-chain rearrangement, and why is TCR gene recombination exclusive to T cells? One possible clue is the observation that susceptibility to recombination seems to be correlated with transcriptional activity of germline gene elements. For example, Reth and Alt (273) reported that AMuLV-transformed pre-B cell lines—representing a developmental stage capable of rearranging VH to DJH—synthesize an RNA transcript that includes DJ and Cμ sequence (termed Dμ RNA). Furthermore, many germline VH genes are transcribed at the pre-B cell stage, just at the time when these genes are targets for recombination (274); these transcripts—designated sterile transcripts because they do not encode a functional Ig chain—are not seen in more mature B cells in which H-chain recombination has been terminated (275). Similar sterile transcripts have been reported for other germline Ig gene elements during the period when they are actively rearranging. Susceptibility of a segment of DNA to both transcription and recombination might be a reflection of a common chromosomal state (accessibility) required for both reactions, or transcription itself might be a prerequisite for recombination, perhaps by partially unwinding the DNA. Interestingly, a similar correlation between transcription and recombination has been reported for the isotype switch of H-chain genes, discussed above, and recombination of yeast mating-type genes.

To further explore the relationship of transcription to recombination, several groups have deleted enhancer regions known to stimulate transcription of the murine κ or H-chain locus and found that recombination of the corresponding locus was substantially reduced. In the κ locus, one enhancer (iEκ) is located in the intron between Jκ and Cκ, and a second enhancer (3′Eκ) is located about 9 kb downstream from Cκ. When homologous recombination in ES cells was used to replace Eκ by a neomycin resistance gene (neo^r), homozygous mutant mice were found to have no κ gene rearrangements (276). It should be noted that another report suggests that replacement by neo^r may impair κ gene recombination more effectively than simple deletion (277). Compared with effects of Eκ elimination, deletion of the 3′Eκ caused a more modest reduction in κ gene rearrangement (278). Similar conclusions on the importance of gene enhancers in supporting recombination have been obtained with transgenic miniloci capable of V(D)J recombination (279) and with similar constructs stably integrated into cell lines competent for V(D)J recombination (280). However, the relationship between transcription and recombination is not simple. One report suggests that in transgenic constructs the 3′κ enhancer actually downregulates recombination (264). And in mouse strains transgenic for another κ minilocus, linkage to a rabbit iEκ substantially increased recombination even though this enhancer is inactive in upregulating transcription in mouse cells (281). Finally, two elements known as KI and KII, which are located just upstream from Jκ1 and have no known enhancer function, appear to be important for Vκ-Jκ recombination because such recombination was substantially inhibited in B cells containing disruptions in these elements, at least under certain conditions (282). Clearly, further investigation will be necessary to clarify how gene recombination is controlled by transcription and chromosomal changes. The further question of how these parameters are themselves regulated is addressed below (with respect to feedback regulation by Ig proteins) and in more detail later in this chapter.

Allelic Exclusion Models

Two general models have been proposed to explain allelic exclusion and isotype exclusion. The stochastic model interprets the observed high frequency of defective rearranged genes as a consequence of the rarity of functional rearrangements; allelic exclusion would follow from the low probability of the coincident occurrence of two rare events in the same cell (283). In this model, the low frequency of λ-producing cells in the mouse would be a stochastic consequence of the smaller repertoire of λ V regions available to rearrange. An alternative to the stochastic model, which might be called the regulated model, was first proposed by Alt and colleagues (284) and has received considerable experimental support. According to this model, the functional rearrangement of an L (or H)-chain gene in a particular B cell would inhibit further L (or H)-chain gene rearrangement in the same cell. If the inhibition occurred promptly after the first functional rearrangement, then two functional Igs could never be produced in the same cell. An initial nonproductive rearrangement would have no inhibitory effect, so recombination could continue until a functional product resulted or until the cell used up all its germline precursors.

Model for Regulated Recombination

As a framework for discussing this model in more detail, Fig. 15 illustrates four hypothetical regulatory influences that could be

components of an allelic exclusion control mechanism. As shown in the upper left corner of the figure, the first Ig gene rearrangements that occur in a B-lineage cells join D to JH segments. The resulting DJ junctions are commonly seen on both chromosomes in early B-lineage cells from fetal liver or bone marrow that have been transformed by AMuLV, as well as in pro-B cells isolated from normal bone marrow by flow cytometry (285). Analyses of cells at this early stage consistently show κ and λ genes in germline configuration. The next recombination step is V→DJ. The expected frequency of VDJ joints maintaining the proper triplet reading frame from V to J is about one third, so most VDJ junctions will be nonfunctional. In addition, some rearranged H-chain genes may be nonfunctional despite in-frame VDJ junctions (286) owing to defects in the germline VH sequence. According to the model, if the initial VDJ junction is nonfunctional for any reason, then in the absence of a viable μ protein, H-chain gene recombination can continue on the other chromosome. If the VDJ recombination on the second chromosome is also nonfunctional, then the cell may have reached a dead end, leading to death by apoptosis (gray shape in Fig. 15). The apoptotic fate of such nonproductive cells has been supported by the observation that mice transgenic for the apoptosis suppression gene *bcl-x*$_L$ harbor an expanded population of bone marrow pro-B cells with almost all nonproductive VDJ joints (9). Although secondary rearrangement may rescue some cells with two nonfunctional recombinations, it is not clear how frequently such secondary events occur.

In contrast, if the first V→DJ recombination in a pro-B cell produces a functional VDJ gene, then its expression will lead to the synthesis of μ H chain. This H chain is expressed in the surface of pre-B cells as a pre-BCR in association with the SLCs VpreB and λ5 (as discussed above) and the same Igα-Igβ heterodimer that, in mature B cells, transmits into the cell the activation signal initiated by antigen-induced cross-linking of surface IgM (see Chapter 7). This pre-BCR complex has two regulatory effects: it blocks further H-chain recombination (① in Fig. 15) and it activates κ gene rearrangement (② in Fig. 15). The μ-induced block to further H-chain rearrangement was initially hypothesized based on the static analysis of myelomas (i.e., as an explanation for observed allelic exclusion), but more recently it has been directly supported by experimental manipulation of Ig genes. When a functionally rearranged μ transgene was inserted into the genome of a mouse strain, it was found to markedly suppress the rearrangement of endogenous H-chain genes in B-lymphocytes (287), suggesting that the protein product of the μ transgene, expressed in pre-B cells, could shut off V(D)J recombination of germline elements in the endogenous IgH locus. A transgene encoding the membrane form of μ (μ_m) was competent to suppress endogenous VDJ recombination, but one encoding only the secreted form (μ_s) was not (288–290), demonstrating that a membrane form of μ protein is required to mediate allelic exclusion. This implication is supported by the observation that allelic exclusion is lost in mouse strains that—due to gene targeting—cannot express the μ membrane exon (291) or functional λ5 protein (292) that is necessary for surface Ig expression; in these animals, individual B cells may carry two productive V(D)J junctions because any μ protein resulting from an initial recombination on one allele cannot assemble on the membrane as a pre-BCR to shut off V→DJ rearrangement of the other allele. The signal for suppressing VDJ recombination appears to be mediated by the Igα-Igβ heterodimer; a μ_m transgene in which critical residues mediating association of μ H chain with this heterodimer were mutated did not suppress endogenous VDJ recombination, but when this transgene was engineered to express the cytoplasmic domain of Igα or Igβ, the resulting chimeric transgenes were able to shut off endogenous VDJ recombination (293,294). The normal pre-BCR–induced shut-off may be mediated in part by downregulation of RAG gene expression (295). This view would be consistent with the levels of RAG gene expression detected in murine bone marrow cells sorted by flow cytometry into populations representing different stages in B-lymphocyte development: RAG-1 and RAG-2 mRNAs were detectable in pro-B and early pre-B cells (corresponding to cells undergoing D→J and V→DJ recombination), but undetectable in the large prolifer-

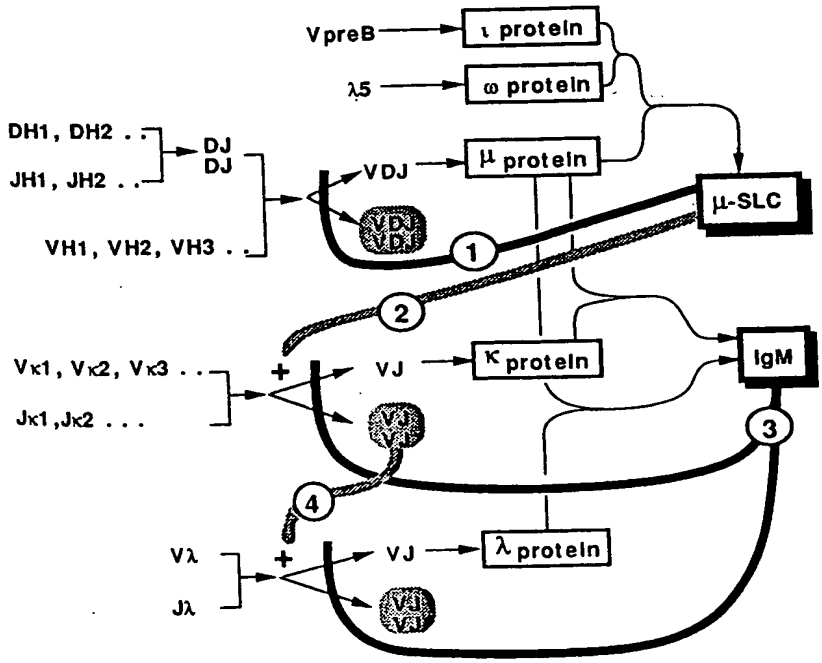


FIG. 15. Regulation of V assembly recombination. Allelic and isotypic exclusion can be explained by the four regulatory effects diagrammed here. Negative regulatory effects (*thick black lines*) prevent a second H- or L-chain recombination event from occurring after an earlier event leads to a functional protein product. Positive regulatory effects (*thick gray lines terminating in +*) switch on κ recombination only after a functional μ protein is produced, and switch on λ recombination only after nonproductive κ recombination has occurred on both chromosomes (*shaded oval area*). The latter effect has not been demonstrated unequivocally; alternatively, a functional μ protein may activate recombination in both κ and λ loci.

ating preB-II cells expressing μ-SLC. RAG gene expression then becomes detectable again in the small preB-II cells undergoing L-chain V→J recombination. In addition to effects on RAG activity, the pre-BCR may also downregulate further VDJ recombination by reducing accessibility of the H-chain locus, as indicated by reduced sterile VH gene transcription (296) and by reduced ability of RAG proteins to produce broken signal ends in nuclei incubated in vitro, as determined by LMPCR (297). Diminished accessibility of the IgH locus would prevent further V→DJ recombination during the subsequent stage when RAG proteins are upregulated to activate L-chain VJ recombination.

Interestingly, the signals mediated by the pre-BCR (μ-SLC) and the mature BCR (IgM) are critical not only for regulating VDJ recombination but also as checkpoints controlling other features of B-lymphocyte differentiation. Thus, in bone marrows of RAG knockout (298) or JH knockout mice (299), B-lymphopoiesis appears to be blocked at the earliest pro-B stage: (a) large cells stain positive for B220, CD43, and c-kit, and (b) cells with surface markers typical of mature B cells are absent from the periphery. When a recombined VDJ-Cμ H-chain transgene was introduced into a RAG-1-/- or RAG-2-/- background (300,301), the resulting μ protein allowed the progression of B-lineage cells to the stage of small preB-II cells, where L-chain recombination would normally occur. These cells could not undergo VL→JL recombination in the absence of RAG proteins but did show upregulation of sterile κ transcription, an apparent reflection of regulatory effect ② in Fig. 15. If, in addition to the μ gene, a complete recombined L-chain transgene was also added to the genome of the RAG knockout mice, then B cell development appeared to be restored, with normal numbers of B cells in the periphery, expressing mature B-cell surface markers and secreting antibody encoded by the transgenes. (An L-chain transgene alone was ineffective in "rescuing" B-cell differentiation.) The developmental block in RAG knockouts is similar to that seen in λ5 knockouts, which lack the SLC component of the pre-BCR. These mice are also arrested in pre-B cell maturation because of the absence of pre-BCR signaling. A similar immunodeficiency syndrome has recently been reported in humans with homozygous defects in the human homolog of λ5 (301a). Interestingly, the λ5 knockout mice could be rescued from their developmental arrest by a recombined κ transgene that was expressed in pre-B cells, indicating that a κ chain can substitute for the SLC in mediating maturation signals (302); indeed, even without the κ transgene, some maturation occurs in the λ5 knockout mice, a presumed result of small amounts of Vκ-Jκ recombination occurring before VDJH recombination and thus providing a κ chain that allows surface IgM expression and signaling. The permissive effect of the pre-BCR on developmental progression appears to be mediated by the Igα-Igβ heterodimer, based on results with the mutant or chimeric μ transgenes linked to Igα or Igβ cytoplasmic domains, as described above (293,294).

The hypothesis that μ protein can activate κ-chain recombination (effect ② in Fig. 15) was originally derived from static comparisons of various B-lymphoid cell lines: cells in which only μ genes are rearranged and expressed are common (pre-B cell lines), but κ-expressing cells without H-chain gene rearrangement and expression are rare, as though H-chain expression were a prerequisite for κ expression. This view has been supported by observation of AMuLV-transformed lines (187) and normal B-cell precursors cultured *in vitro*: these lines always rearrange H-chain genes before κ genes. In a more direct demonstration that μ protein could stimulate κ rearrangement, an AMuLV-transformed cell that could not express endogenous H chain (because of defective VDJ rearrangements on both of its H-chain loci) was found to retain the κ locus stably in germline configuration until a functional μ gene was introduced; this μ gene activated κ gene rearrangement and expression (189,303). Only a gene encoding μ_m was effective in activating κ rearrangement and not a μ_s gene, again suggesting a requirement of surface pre-BCR expression for this signaling. A similar conclusion was deduced from the rescue of RAG knockout animals by a recombined VDJ-Cμ transgene, which caused upregulation of sterile κ transcription in preB-II cells, as described above. Also, a human μ transgene was found to upregulate both κ gene (sterile) transcription and Vκ-Jκ rearrangement in B-lymphoid precursors in fetal liver (296). To directly assay for accessibility of Ig loci to RAG proteins, Constantinescu et al. (297) incubated RAG-/- nuclei from different B-cell developmental stages with RAG proteins in vitro and detected broken signal ends by LMPCR; using this method, they found that introduction of a VDJ-Cμ transgene into the RAG-/- background caused a 30-fold increase in the frequency of breaks at Jκ1 observed in pre-B cell nuclei incubated with RAG proteins. Interestingly, a δ transgene is also apparently able to both inhibit endogenous VDJ rearrangement and activate κ rearrangement (304), so these effects must be mediated by properties common to δ and μ H chains. Crosslinking of the pre-BCR complex is apparently required for activation of κ recombination (163), suggesting that a ligand that occurs physiologically in the environment of the pre-B cell may bind to the pre-BCR to signal activation of gene rearrangement in the κ locus and suppression in the H-chain locus; however, no candidate ligand has been identified.

Despite all the evidence cited above, it is clear that some sterile κ gene transcription Vκ-Jκ recombination can occur in the absence of a μ_m-containing pre-BCR (291,305–307). This may be a consequence of some "leakiness" of the controls on L-chain recombination in early B-lymphopoiesis or may reflect a separate developmental lineage in which Vκ-Jκ recombination is activated earlier; but the low frequency of this premature Vκ-Jκ recombination would keep the possibility of L-chain double-producers violating allelic exclusion below 1% (299).

When κ recombination begins, the possibilities for functional and nonfunctional VJ rearrangements resemble those discussed above for the H chain. According to the regulatory model, if a cell initially rearranges its κ locus on one chromosome nonproductively, then it can proceed to rearrange the locus on the homologous chromosome. As soon as functional κ gene rearrangement leads to expression of a functional κ chain that can associate with μ to form a surface-expressed IgM molecule (i.e., a mature BCR), then further κ rearrangement will be suppressed (effect ③ in Fig. 15). This regulatory influence would explain the observation of allelic exclusion in κ-expressing myelomas, and it has been supported by the finding that a functional rearranged VJ-Cκ transgene can suppress rearrangement of endogenous κ genes (308). Furthermore, in murine B-lymphoma lines expressing RAG proteins, cross-linking of surface IgM with an anti-μ antibody was found to rapidly suppress RAG gene expression (309). A ligand might deliver a corresponding signal in physiologic circumstances, but this is not clear, especially because under some conditions cross-linking the BCR of pre-B cells (as might occur on binding of a self-antigen) can actually upregulate RAG gene expression to activate "receptor editing." A BCR-mediated signal seems to be required for small preB-II cells to advance to the immature B phenotype and move into the periphery because, as mentioned above, RAG knockout animals rescued with only a μ transgene were not able to advance beyond the pre-B stage, whereas a combination of μ and κ transgenes allowed normal B-cell proliferation, surface marker expression, and migra-

tion to the periphery. Igα is also apparently critical for mediating some differentiation signals of the BCR because mice with an engineered deletion in the cytoplasmic signaling domain of Igα had only 1% of the normal numbers of circulating B cells (310).

Regulatory effect ④ in Fig. 15 is rather speculative because little is known about regulation of λ gene rearrangement. The idea that λ recombination is somehow triggered by nonfunctional κ rearrangements on both chromosomes derives from the observations that most B cells show isotypic exclusion (i.e., express either κ or λ but not both) and that κ rearrangement seems to occur before λ. Thus, in studies of normal and malignant human B-lymphoid cells (136,311), in κ-expressing cells, λ genes were found to be in germline configuration, whereas in λ-expressing cells, κ genes were either rearranged (rarely) or deleted (most commonly). The κ deletions reflect the RS recombination event discussed earlier in this chapter. These results suggest that λ genes remain unrearranged until κ genes rearrange nonproductively or are deleted. The mechanism of this apparent regulation of the λ genes by the κ locus is unknown. It has been suggested that somewhere in the 24 kb between Cκ and the RS site lies a sequence that suppresses λ rearrangement but is deleted in the κ RS recombination event to alleviate this suppression. However, in contrast to this hypothesis are the observations that mice with a targeted deletion of either Cκ or the intronic κ enhancer have almost normal numbers of B cells, essentially all of which express λ despite having no loss of DNA between C_κ and Rs (276,306,312).

It is assumed that membrane expression of a μ-λ IgM would shut off all further L-chain gene recombination by a similar mechanism to that in the κ locus (as illustrated by effect ③ in Fig. 15), a supposition supported by suppression of κ gene expression in λ transgenic mice (313). However, this suppression is somewhat leaky, especially in older mice (314–317), and even in normal splenocytes a small population of cells expresses both isotypes (318). These observations have led to the speculation that certain B-lymphocytes are not programmed for strict isotype exclusion. It is also possible that in some cells expressing both κ and λ, one of the isotype L chains has such a greater affinity for the expressed H-chain protein (on the basis of VH-VL compatibility) that the other isotype does not functionally contribute to surface Ig and is thus allelically excluded at the protein level. Some evidence in opposition to the strictly ordered (κ before λ) rearrangement model for L chains has been put forward (285,319), suggesting that VJ recombination is activated concurrently in both the κ and λ loci. In this model the preponderance of κ-expressing lymphocytes (in mice at least) would result from a stochastic process in which κ rearrangements are favored by the larger Vκ repertoire and other features of the κ locus, such as more active recombination signal sequences (320). Alternatively, the preponderance of κ expression may be explained by a model in which κ and λ rearrangement occurs independently, but κ recombination initiates earlier in B-cell development (321). In either of the latter two models of L-chain rearrangement, a surface IgM molecule ($H_2\kappa_2$ or $H_2\lambda_2$) would signal feedback suppression of L-chain recombination, so that both allelic exclusion and isotype exclusion would be explained by the same mechanism (and the regulatory effect ④ of Fig. 15 would not exist).

RAG Protein Production after Mature BCR Expression

Although the RAG genes are apparently downregulated through a signal mediated by the appearance of IgM on an immature B cell, there is evidence that RAG gene expression occurs in at least two later stages of B-cell development: during receptor editing of autoreactive B cells in the bone marrow and during B-cell maturation in germinal centers.

An early observation suggesting the possibility of receptor editing by secondary rearrangement of κ L-chain genes came from an analysis of circular DNAs representing the deleted segment in VJ recombination (322). It was observed that, in addition to containing the expected signal joints, many of these circles contained VJ junctions; these could have formed in an initial inversional recombination, which was then followed by a secondary deletional rearrangement that produced the observed circular DNA. Significantly, about a third of the VJ junctions analyzed showed no apparent defect, indicating that they could have produced a functional antibody that was altered by secondary recombination. Since this study, several laboratories have directly observed secondary κ rearrangements in B-cell tumors and AMuLV-transformed pre-B lines (188,323). H-chain V-gene replacement also can occur, mediated by a 7-mer embedded in the 3′ end of many VH coding regions, as described in an earlier section. As in the case of κ genes, such replacement can occur even after a productive VDJ recombination (180,324). A potential reason for replacing productively rearranged L or H chains would be to abort production of an antibody that was autoreactive. Thus, receptor editing might complement two other mechanisms for preventing autoantibodies: anergization and cell deletion by apoptosis.

Several studies have supported this interpretation using mice carrying transgenes expressing autoreactive antibodies (325,326). In one study, the JH locus was targeted for replacement by homologous recombination with the 3H9 recombined VDJ gene; this gene encodes an H chain that in combination with most (but not all) κ L chains can bind to DNA, a self antigen (181). In such mice, most B cells have replaced the 3H9 gene by an upstream VH gene, with junctions showing typical N regions and exonuclease nibbling. When inserted as a normal transgene, 3H9 also stimulates L-chain editing, as evidenced by increased frequency of Jκ5 usage and reduced diversity of Vκ genes expressed by the B cells displaying the 3H9 H chain (325). These results are consistent with the interpretation that primary rearrangements, yielding Vκ proteins capable of supporting DNA binding, were edited by secondary rearrangements involving downstream Jκs and Vκ regions incompatible with DNA binding. Receptor editing appears to occur in the immature B-cell population in the bone marrow (327,328) and is associated with increased RAG gene expression. Indeed, BCR cross-linking of a human B-cell line was found to upregulate RAG gene expression (329). As previously discussed, BCR cross-linking also has been reported to terminate RAG gene expression of surface IgM^+ immature B cells to mediate allelic exclusion (309). How a cell can discriminate between a BCR-mediated signal that downregulates RAG expression to mediate allelic exclusion and a BCR-mediated signal that upregulates RAG expression to initiate receptor editing is currently not understood; but differences in receptor affinity, the precise stage of development, or costimulatory signals might be involved. Interestingly, a failure of receptor editing may contribute to the autoantibodies mediating systemic lupus erythematosis (330).

A second instance of late RAG expression occurs in germinal center (GC) B cells (331,332). RAG-1 and RAG-2 mRNA transcripts were detected by RT-PCR in FACS-purified GC cells from immunized mice, the RAG proteins were detected in GC cells by immunofluorescence, and evidence of ongoing V(D)J recombination was found in GC cells (332a,332b). RAG expression also was observed constitutively in Peyer's patch GCs (which are maintained by food antigens in the absence of intentional immunization) and

in splenic B cells cultured with IL-4 plus LPS (conditions known to induce at least one other process typical of GC cells, i.e., isotype switching). GC cells appear to recapitulate expression of several surface markers characteristic of early B-lineage cells, including heat stable antigen (CD24) and λ5; so it appears that the RAG gene expression may be just one aspect of a GC-induced reversion to a primitive phenotype. What function might RAG proteins have in the GC? The primary processes that affect Ig expression in the GC are somatic mutation and isotype switching, but it is unlikely that RAG proteins are expressed in GC to function in either of these processes because both can occur in B cells of RAG-/- mice (116a,331). One possibility is that RAG-dependent receptor editing may be turned on to replace V regions that have become autoreactive as a result of GC-induced somatic mutations; further experiments will be necessary to evaluate this interpretation.

GENERATION OF DIVERSITY

One of the most interesting questions about Igs is the source of the immense variation observed in antibody-binding specificities. As discussed at the beginning of this chapter, early speculative debates about this question centered on the relative contributions of germline repertoire and somatic mutation in creating diversity. One source of diversity that was unanticipated before the recombinant DNA revolution already has been discussed in this chapter in some detail: somatic recombinational diversity. We will now focus on the germline repertoire and then consider the contribution of somatic mutation.

Germline Diversity

A comprehensive evaluation of the germline repertoire of V-gene segments requires an examination of the sequences of all germline V regions, a daunting task. However, modern molecular biology techniques—including cloning vectors allowing long genomic inserts and large-scale sequencing with fluorescent dyes and automated sample preparation—have helped realize this goal for the human κ, λ, and H-chain loci; and considerable progress has been made with the murine κ and H-chain loci. (The tiny V repertoire of the murine λ loci has already been discussed in the section on λ genes.)

Two Worldwide Web resources are devoted to providing convenient updated access to Ig germline gene sequences. The IMGT (international ImMunoGeneTics) data base (http://imgt.cnusc.fr:8104/home.html), coordinated by Marie-Paule Lefranc, includes a data base for Ig and TCR sequences, as well as a separate one for major histocompatibility complex (MHC) sequences (333). In the Ig/TCR database, all species for which data are available are included; sequences are annotated in standard formats, and map information is provided graphically. V Base Gold (http://www.mrc-cpe.cam.ac.uk/imt-doc/public/INTRO.html) is an online catalog of human V gene segments and alleles coordinated by Ian M. Tomlinson.

Germline Diversity of the Murine IgH Locus

Attempts to analyze the mouse VH repertoire began before the gene cloning era with the study of VH amino acid sequences from mouse myelomas. Initial attempts to classify the observed VHs into related groups were based on limited amino acid sequence analysis, primarily of N-termini of myeloma proteins. The current scheme classifies two V gene sequences into the same group or family if they show more than about 80% nucleotide sequence identity, and into different families if their sequences are less than 70% identical. (Empirically, few VH comparisons yield identities between 70% and 80%.) These criteria for sequence similarity correspond well with the degree of similarity that allows hybridization between a V probe and members of the same family under conditions of moderate stringency. The initial classifications based on this scheme identified seven VH families (334,335). Since that time, continuing analysis of new sequences has identified eight additional families (336–338). The families now known contribute to the bulk of the immune response: when 2,000 cDNA clones hybridizing to both Cμ and JH probes were analyzed, all had V regions from the 15 families, based on hybridization or sequence analysis, except for about 2% of the clones representing truncated or aberrant cDNA synthesis (338). The families have been further classified into three groups, or clans, based on sequence conservation in the framework I region (FR1; codons 6–24) and FR2 (codons 67–85) (339–341). (Framework amino acids are the non-CDR parts of the Ig V region that hold the CDR loops in position to contact antigens.) The clans are conserved between humans, mice, and frogs, suggesting that several fundamental steps in germline VH diversification preceded the amphibian-reptile divergence (342).

The classification of VH genes into families leads to two approachable questions: (a) How many genes in each family are available to contribute to Ig diversity? (b) How are the families arranged on the chromosome? One straightforward approach to the question of gene number is to count the number of bands visible on Southern blots. However, this method can yield only a rough estimate of gene number because of several complications in the interpretation. The number of bands may underestimate the number of VH regions for two main reasons: (a) a given DNA fragment may contain more than one V-related sequence; and (b) some observed bands may actually represent several comigrating hybridizing DNA fragments, each containing different VH genes. On the other hand, the number of bands could theoretically exceed the number of different VH genes for two reasons. First, some hybridizing sequences may not contribute effectively to sequence diversity. In particular, some nonallelic gene pairs may be so similar that the second copy provides no gain in amino acid sequence diversity. Other V sequences are nonfunctional because they have become separated from the C-region locus, even lying on a different chromosome. Still other germline clones isolated on the basis of hybridization to a V probe have turned out to contain multiple defects that would preclude their expression as functional V regions even if they underwent rearrangement (i.e., they are pseudogenes); conceivably these pseudogenes could contribute to diversity at the somatic level through gene conversion, a mechanism known to operate in chickens and rabbits as discussed later. Second, some bands could be counted twice because of hybridization to probes of two different families. This could occur if a specific DNA fragment carried germline V sequences from two different families or if hybridization occurred across group boundaries owing to a clustering of residues identical to the probe sequence.

With these caveats, Table 1, modified from a compilation by Kofler et al. (343), is presented to give an idea of the widely varying complexity of the different VH groups. Several groups have only a few members. For example the VH S107 family yields four Southern blot bands, and extensive cloning with a probe for the family has in fact detected only four germline members (of which one is apparently a pseudogene). At the other extreme, the VH J558 family shows much greater Southern blot complexity and may contain as many as 1,000 members in the BALB/c mouse. This estimate (344) was based on quantitative kinetics of hybridizatic

using an excess of single-stranded J558 probe, whereas the lower estimate of 60 genes is based on counting Southern blot bands. If the larger estimate is accurate, then many of the bands observed on Southern blots with the VH-J558 probe must contain multiple comigrating DNA fragments; these would probably represent recent duplications in the VH locus and would be expected to encode minimally diverged VH sequences. Other strains of mice besides BALB/c seem to have smaller J558 families (345), consistent with the notion of a recent expansion of J558 VH genes in BALB/c.

The question of how the germline VH genes are arrayed on the chromosome has been approached by several different techniques. One straightforward method has been to screen phage libraries of germline DNA with VH probes and to examine clones containing more than one VH region. Application of this method to the mouse VH locus has yielded three important generalizations: adjacent VH genes are usually members of the same family; they are oriented in the same 5′-3′ direction; and they are spaced about 7 to 15 kb apart (346–348). The first finding suggests that members of a given VH family are clustered together on the chromosome. Such clustering represents a simplifying principle that allows the mapping of murine VH genes to be conceptually divided into establishing the order of the family clusters on the chromosome and then establishing the order of VH genes within each cluster.

One approach to ordering the VH families has been to examine (by Southern blotting) the VH bands that are deleted in various myelomas or hybridomas. Deductions about the order of VH families can be made if the nonexpressed chromosome is deleted so that all the VH fragments observed on a Southern blot can be considered to derive from the same (i.e., the expressed) chromosome and to lie upstream from the rearranged VH gene (349–351). A powerful variant of the deletion method has used a panel of Abelson virus–transformed B-cell lines constructed from F1 animals heterozygous for allotype at the IgH locus. In most of these cell lines it was possible to distinguish deletions on the two parental chromosomes and establish an independent VH gene order for each (338,352).

Another mapping approach is brute force chromosome walking by the generation of large overlapping clones using cosmid libraries; this approach is being pursued by a number of laboratories, but is made difficult by the occurrence of recent duplications leaving nearly identical DNA segments that cannot easily be distinguished or ordered on a map. This problem should hopefully be resolved by the use of yeast artificial chromosome (YAC) clones, which can accommodate 1- to 2-Mb segments of genomic DNA. YACs have been useful in long-range mapping of the human V loci. Pulsed field gel electrophoresis also has been used to separate large fragments of genomic DNA for mapping by Southern blotting.

Studies using these techniques have been in general agreement about the order of certain murine VH families, but a complete map consistent with all data from the various mapping methods is not currently available. It is clear from several laboratories that some interdigitation between families occurs (351,353), and this could contribute to difficulties in interpretation. A representative map of 15 murine VH families is shown in Fig. 16, based on the work of Brodeur and colleagues (352,354).

Among VH maps based on different techniques, the best agreement is on the families closest to Cμ. The map order of these families is: S107—Q52—7183—D—J—Cμ, with some overlap between these three families as shown in the Fig. 16. This order is of special interest because the most proximal family cluster (designated 7183), and in particular its most proximal member (designated VH81X), is the V region that is significantly over-represented in the VDJ rearrangements occurring in fetal liver pre-B cells (355). This observation was earlier taken as evidence favoring a tracking model of V gene rearrangement [i.e., a recombinase would engage DNA near the J regions and slide 5′ to find V regions to recombine (356)]; but alternative interpretations have been proposed based on more recent data (339,340,357,358).

Mouse Germline DH and JH Regions

D regions were initially hypothesized based on the highly diverse amino acid sequences in myeloma proteins between the V

TABLE 1. *VH region families of mice and humans*

Mouse					Human	
Family number	Family name	Complexity[a]	Clan[b]	Group[c]	Family number	Complexity[d]
VH2	Q52	15				
VH3	36–60	5–8	II	I	VH2	4
VH8	3609	7–10			VH4	9
VH12	CH27	1			VH6	1
VH1	J558	6–1000				
VH9	VGAM3-8	5–7	I	II	VH1	14
VH14	SM7	3–4			VH7	5
VH15	VH15	2			VH7	5
VH4	X-24	2				
VH5	7183	12				
VH6	J606	10–12	III	III	VH3	46
VH7	S107	3				
VH11	CP3	1–6				
VH13	3609N	1				

This table is based on a compilation of murine VH regions by Kofler et al. (343), and a review of human VH regions by Pascual and Capra (822) and a paper by Mainville et al. (338). Original references for the data can be found in those sources.

[a]Complexity, an estimate of the number of VH sequences in each family.
[b]Clans of VH sequences as defined by Schroeder et al. (340)
[c]Groups, based on the classification by Tutter and Riblet (339).
[d]Complexity, based on the prototypic haplotype provided by Cook and Tomlinson (381).

and J regions, as briefly discussed earlier in this chapter. Because both VH and JH were known to be flanked by signal elements of the long space type (23 bp), it was predicted that a germline D region would be flanked on both sides by short signal element spaces so that both V-D and D-J recombination would conform to the 12/23 spacer rule (359).

Finding a germline D gene with a probe corresponding to D-region sequence from one of the cloned recombined genes was technically difficult because the DNA segment encoding the few amino acids of the D region would be too short to give a usable hybridization signal. To obtain more effective probes, DNA fragments have been isolated from DJH intermediates that have not yet recombined with a VH gene, and thus retain the 5′ flanking sequences from the germline D providing a longer probe (360). A DJH intermediate cloned from the myeloma QUPC52 identified its germline D precursor—designated DQ52—0.7 kb 5′ to JH1. Its structure was very similar to expectation: a 10-nucleotide coding sequence flanked on both sides by RSS elements with a 12-bp spacing. A similar clone derived from a rearranged DJH in a T-cell line (SP2) was used as a second probe to clone nine related D regions clustered within a 60-kb region, all having 17-nucleotide coding segments and short spacing of the signal elements (361,362). A third D probe called FL16, derived similarly, identified the FL16 family, which is composed of only two germline D genes but is well represented in the rearranged IgH-chain genes that have been sequenced. Finally, a last D region, D_{ST4}, was identified through the recognition of a recurring nucleotide sequence observed between V and J in recombined VDJ regions that was not accounted for by the previously known D sequences (363). The 13 murine D regions span about 80 kb upstream from the four JH segments that in turn lie upstream from Cμ.

Apart from the additional combinatorial diversity contributed by the repertoire of germline D elements, the flexibility of the recombination site applies on both ends of the D region. Furthermore, an out-of-frame recombination at the VD junction may be compensated by the frame of the DJ junction so that a particular D element could theoretically be read in all three frames in different VDJ recombinants. As mentioned previously, this extra source of diversity is used by human H chains (11), but the murine system has evolved mechanisms that strongly favor the reading frame known as RF1 (10). DJ rearrangement in RF3 is counterselected owing to frequent internal stop codons. When DJ recombination has occurred in RF2, the resulting transcripts can encode a DJ-Cμ protein (designated the Dμ protein), which can be expressed on the surface of a pre-B cell in association with the products of the VpreB and λ5 genes (364–366). Murine cells expressing Dμ protein cannot progress to normal Ig production, perhaps because the Dμ protein triggers the shut-off of VDJ recombination before V assembly is complete; therefore, expressed H-chain V regions rarely include a DJ junction in RF2 (10). This curious model is supported by the observations that RF2 suppression is not observed in λ5 knockout mice (which fail to express Dμ protein on the cell surface) (367) and that analysis of recombination in single cells by PCR failed to detect cells containing both a DJ junction in RF2 as well as a productive VDJ junction (368). In humans this mechanism is not operative because ATG initiation codons are not generally present 5′ from D regions to encode a Dμ protein. Some rearranged VDJ sequences seem to be interpretable as V-D-D-J products, even in cases where D-D recombination would seem to violate the 12/23 rule (369).

Murine Germline Vk Locus

Although murine L-chain genes were among the first Ig genes studied by molecular biology techniques, the organization of the mouse germline Vκ locus has been less thoroughly characterized than the human. On the basis of N-terminal amino acid sequence data, Potter and colleagues classified mouse Vκ sequences into 24 groups (370). Current classification based on the nucleotide sequence criteria described above recognizes about 20 families (371). However, different classification schemes have yielded different estimates, perhaps because the Vκ genes show degrees of relatedness that are not discrete steps, but rather continuous gradations (372), as would be expected if gene duplications could occur on a time continuum and rates of sequence diversification could also vary. As described above for VH genes, some Vκ families are

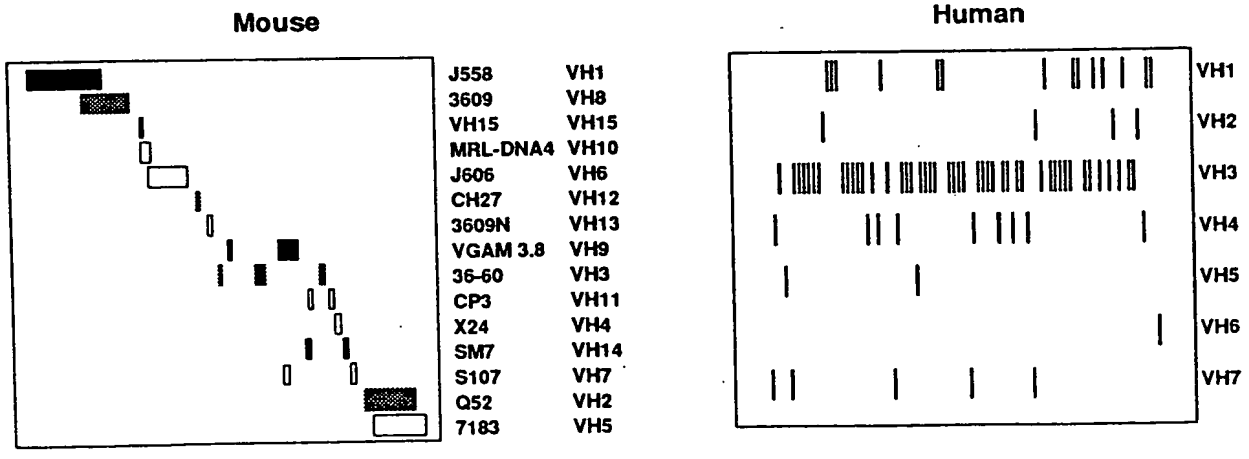


FIG. 16. Maps of the murine and human VH loci. The 15 known murine VH gene families are shown in their approximate map positions. Each rectangle represents a cluster of VH genes of the indicated family; the clan identification (340) of the VH families is indicated by the color of the rectangle: black for clan I, gray for clan II, and white for clan III. Although some interdigitation is shown by overlapping families (e.g., the Q52 and 7183 families), the families are largely clustered. In contrast, all human VH genes (*vertical lines*) of a prototypic haplotype are shown in the right panel, based on the formulation by Cook and Tomlinson (381); extensive interdigitation of families is apparent.

shared by humans and mice, suggesting that the family divisions preceded primate–rodent species divergence (373,374). The murine locus, including about 140 Vκ sequences (genes and pseudogenes), has been cloned on a series of overlapping bacterial artificial chromosome (BAC) and YAC clones, and spans about 3.5 Mb upstream from the Cκ gene on chromosome 6 (375–377). In addition, a few Vκ sequences have been localized to other chromosomes (chromosome 16 and 19), where they could not contribute to diversity and are thus considered "orphons." In the functional Vκ locus on chromosome 6, many related Vκ sequences are found to lie clustered together, although some interspersion of families also exists. As of this writing, the total number of genes (versus pseudogenes) has not been completely determined by sequence analysis.

Human Germline VH Locus

Amino acid sequences of human myeloma VH-region proteins were originally classified by Kabat into three groups. These have been found to correlate well with three families defined by sorting together VH probes that produce similar Southern blot patterns. Four additional human VH families have been found more recently using molecular genetics approaches. Several of the human VH families show sequence similarities to particular mouse VH families (378), and all can be classified within the three large clans of murine VH genes common to the mouse, human, and Xenopus; as noted above, these observations suggest that significant germline VH diversification antedated the amphibian–reptile species divergence.

Early phage and cosmid clones of human VH genes demonstrated that the human VH families are extensively interdigitated, in contrast to the family clusters characteristic of the murine locus. This interdigitated structure was confirmed by gene mapping studies involving analysis of VH deletion in B-cell lines and of PFGE-based mapping data (354,379); ultimately the complete delineation of the human VH region has been achieved through analysis of overlapping YAC clones covering the entire locus (380–382). The VH locus spans 1.1 Mb at the telomeric end of chromosome 14 and includes 95 VH sequences; of these, 51 are functional and most of the remainder appear to be pseudogenes (although the exact numbers are somewhat variable depending on the haplotype, and a few V sequences have not yet been fully characterized). Each VH region is identified by a two-digit number: the first number corresponds to the VH family and the second indicates the sequential number of the VH on the standard map (starting from the JH proximal end), with the letter P appended for pseudogenes. A particular VH locus is judged to be functional if it has no obvious defects in coding sequence and has been detected in a rearranged VDJ gene, indicating intact recombination signals. Additional polymorphic V regions are designated (using a decimal point) with reference to the JH-proximal standard VH; e.g. a polymorphic V region from the VH7 family lying between 4-4 and 2-5 is identified as 7-4.1. The entire VH locus thus extends from V 6-1, which is located about 77 kb 5′ from JH1 (383), through V 7-81, which appears to be located within a few kb of the telomeric repeat sequences marking the end of the q arm of chromosome 14. Twenty-four additional germline VH sequences have been mapped to chromosomes 15 and 16 and represent nonfunctional orphons that were apparently duplicated from the functional locus on chromosome 14 (384,385); these sequences contributed to earlier overestimates of the length of the functional human VH locus. All the VH regions whose transcriptional orientations have been determined share the same orientation characteristic of the JH regions, consistent with recombination by deletion rather than inversion. Several regions in the locus show evidence of ancient duplications: segments in which a pattern of hybridization to different probes from nearby DNA regions is repeated elsewhere in the locus. Thus, the complete locus map may offer clues to its evolution, as well as defining the repertoire of germline VH diversity available to the immune system.

Human JH and DH Regions

Upstream of the human Cμ gene lies a set of JH-region genes, including six apparently functional JH regions (386). Interspersed among the active human JH genes are three J pseudogenes that encode amino acid sequences never found in human H chains and that lack the RNA splice signal found at the 3′ end of all active JH genes. All of the JH genes and pseudogenes demonstrate 23-bp RSS spacing (as in the mouse).

Complete sequence analysis of a 92-kb region spanning the human D regions (11) has confirmed the general structure of the locus previously deduced from partial sequence analysis and Southern blotting. One germline D gene is located in a position roughly homologous to that of the mouse DQ52, that is, 5′ to the human JHI. This human D gene, initially designated DHQ52, bears striking homology to its murine counterpart but is the only human D segment showing such human/mouse homology. All of the other human D regions fall into six families and lie in a cluster of duplicated domains about 22 kb upstream from JH1. There are 27 D regions; 24 of these are accounted for by four tandem approximate duplications of a 9.5-kb segment containing a representative of the six D families. In addition to these 24 D regions, three more D regions result from (a) an additional partial duplication of 2.8 kb, including one D; (b) an internal duplication creating one D; and (c) DHQ52, which is in a family of its own, distinct from the six duplicated families. The D regions have been renamed following a scheme similar to that used for the VH genes: a first number identifies the family, and a second identifies the sequential position in the locus. The locus starts with the 5′-most D region, D1-1, and ends with D7-27 (DHQ52). Three D regions are apparently nonfunctional as a result of mutations in RSS 7-mers, and there are two pairs of D regions with identical coding sequences (including one of the D segments with a 7-mer mutation); so there are 23 distinct D regions that can contribute to human Ig diversity. A comprehensive computer analysis of a data base of published human VH sequences showed that all of these sequences appear in the data base, many in all three reading frames. In general, one reading frame encodes primarily hydrophilic residues, one encodes hydrophobic residues, and one includes frequent stop codons. (Some D regions that contain stop codons can be used if these codons are removed by nuclease trimming before VDJ assembly is complete.)

In addition to these D regions, all flanked by signal elements with the typical 12-bp spacing, one putative family, designated DIR (D with irregular spacing), has been described, having RSS elements that could be taken with either 12- or 23-bp spacing; theoretically, DIR regions could contribute extra diversity in the form of V-D-DIR-J or V-DIR-D-J rearrangements without violating the 12/23 rule (387). However, the systematic evaluation of the data base of 893 published VH regions failed to detect expression of DIR regions (11) confirming the failure to detect such rearrangements using a sensitive PCR assay (388). Also absent from this VH data base were inverted D regions and the previously hypothesized D-D rearrangements (which would violate the 12/23 rule), although evidence that both of these can occur at low frequency has been obtained using highly sensitive PCR techniques (388,389).

Furthermore, DIR rearrangements, both direct and inverted, were found in mice transgenic for a human IgH minilocus (389a). Additional human D segments originally thought to lie upstream from the main cluster apparently lie on the duplicated orphon cluster on chromosome 15 and are thus nonfunctional (385,390).

Human Germline Vκ Locus

The human Vκ locus is located on the short arm of chromosome 2 (2p11-2). Most of its genes fall into four known V gene families: VκI, VκII, VκIII, and VκIV. Five cloned Vκ sequences have been described that would fall into three additional families (VκV through VκVII), but these sequences have apparently not contributed to known proteins and thus are probably pseudogenes. Zachau and colleagues have performed an extensive investigation of the human locus by cloning, PFGE, and sequence analysis (391,392). They have identified 76 Vκ sequences in the human Vκ locus, lying in two 0.4-Mb contigs separated by a spacer of about 0.8 Mb that is apparently devoid of Vκ sequences. The Jκ distal (upstream) segment, including 36 Vκ sequences, appears to be the result of a large duplication. Within each duplicated segment, all V regions have the same 5′-3′ orientation. Remarkably, the segment distal to the Jκ-Cκ region lies in inverted orientation with respect to the proximal segment and Jκ-Cκ. Most of the duplicated Vκ sequences could be assigned to the proximal or distal segment by preparative separation of the two loci using PFGE, followed by Southern blots exploiting the rare differences in the restriction maps of the duplicated segments. Alternatively, assignments could be deduced from Southern blot bands absent from the DNA of rare individuals lacking the distal duplication. In all B-lymphoid cell lines examined, those with rearrangements involving the distal inverted Vκ segments contained retained signal joints and failed to show deletions of downstream Vκ and 5′ J-flanking DNA, consistent with Vκ-Jκ recombination by inversion. Except for two insertion/deletion differences leading to one unpaired proximal Vκ and one unpaired distal Vκ, the Vκ sequences of the proximal and distal parts of the locus match their homologs with 95% to 99% sequence identity. This high degree of similarity suggests a recent origin for the duplication, which is supported by the fact that the duplication is not found in chimpanzees or gorillas (393), which are thought to have diverged from the human lineage only 6 to 8 million years ago. Between the proximal duplication and the Jκ regions lie an additional six unpaired Vκ sequences, of which the two nearest Jκ lie in inverted orientation. The most J-proximal Vκ sequence, the single gene of the VκIV family, is only 23 kb upstream from Jκ1. Members of the VκI, VκII, and VκIII family have been found to be extensively interspersed. Of the 76 Vκ sequences in the locus, 33 are without apparent defect, although some in the duplicated segments are so similar to their duplicated counterpart that they do not contribute significantly to diversity of the locus, and some may not be expressed, for unknown reasons. In an examination of 70 cDNAs from a human spleen library plus 170 cDNAs from the literature, only 21 of the Vκ genes plus five from duplicated identical genes were found to be expressed, for an expressed cDNA repertoire of 27 Vκ genes (394). Of the remaining Vκ sequences in the germline Vκ locus, 25 are unequivocal pseudogenes, demonstrating several crippling defects; in addition, 16 sequences have one or two minor defects and might be functional in some haplotypes.

Apart from the Vκ sequences in the cluster near the Jκ-Cκ locus, Zachau and colleagues have identified at least 25 orphons. One orphon cluster is located in the long arm of chromosome 2; perhaps it was separated from the major locus—on the short arm of thi chromosome—by a pericentric inversion [which must hav occurred rather recently in evolution because it is absent fror chimpanzees and gorillas (395)]. Other orphons are located o chromosomes 1 and 22; and at least one probably nonfunctional V lies about 1.5 Mb downstream from Cκ (396).

Human Germline Vλ Locus

For many years the human Vλ system was the least characterize of the V loci of human and mouse, but the relative obscurity of thi locus has been dramatically reversed by recent intensive cloning sequencing, and mapping of Vλ regions (168,397) and ultimately b the complete sequence analysis of 1,025,415 bp covering the enti locus (149). The locus contains about 36 potentially functional V genes (in 10 families), 33 pseudogenes, and 34 relics, containin Vλ-like sequences severely disrupted by insertions or deletions. (A noted for other loci, exact numbers may differ depending on th haplotype and method of analysis.) Of the potentially function genes, only about 30 have been documented to be expressed b comparison with cDNA sequences. Within the clustered V sequences lies the human VpreB gene, as well as several genes an pseudogenes unrelated to the λ system. All the Vλ sequences are i the same transcriptional orientation as the J-C cluster. Analysis o the 1-Mb sequence shows several segments of internal duplication some including Vλ regions. The largest and most frequent expressed Vλ gene families lie relatively close to the J-C cluste mostly within the proximal 400 kb. Interestingly, these families a most similar in sequence to the Vλ genes of species that expre predominantly this isotype of L chain, including chicken, horse, ar sheep, whereas the Vλ genes of the BALB/c mouse are most sim lar to the least frequently expressed human families.

Combinatorial Diversity Estimates

Before the era of recombinant DNA technology, the source o antibody diversity was so mysterious that it was whimsical referred to as the problem of generation of diversity (GOD Knowledge of antibody genes gained over the past 20 years has el cidated the diversity inherent in the germline V repertoire plus th diversity contributed by recombinational mechanisms (combinat rial multiplication, flexibility of recombination site, N and nucleotides), as already discussed. Together these diversity el ments provide an immense potential repertoire, one so large that some investigators it seemed unnecessary to postulate that diversi was further increased by somatic mutation. As an exercise in est mating the contribution of germline and recombinational diversi in the human, consider the number of different antibodies th could be formed assuming 39 functional VH genes, 27 Vκ gene and 30 Vλ genes. For κ sequences, we can multiply 27 (Vκ gene × 5 (Jκ regions) × 2 (a conservative multiplier reflecting variabili around residue 96 resulting from flexible recombination), yieldi the product 270. For λ sequences, we can multiply 30 (Vλ genes) 4 (Jλ regions) × 2 (flexibility multiplier), yielding the product 48 Thus, the total VL possibilities are 270 + 480 = 750. For V sequences we have 39 functional germline genes × 23 (DH se ments) × 4 (JH) × 4 (flexibility multiplier on both sides of the segment) × 3 (possible reading frames of the D region) = 43,05 Assuming random association of L and H chains to form a cor plete L_2H_2 antibody molecule, the number of different combin tions is 750 × 43,056 = 32 million. This estimate has neglect additional sources of diversity that are substantial but difficult

quantitate: the insertion of N and P nucleotides. However, even neglecting these factors the exercise demonstrates how nature has greatly enlarged the potential sequence diversity available from a limited number of total nucleotides by allowing flexible recombination between different sequence elements.

Although it is clear that the above mechanisms imply a vast repertoire, it is worth considering some qualifications that tend to reduce the actual combinatorial diversity, especially early in ontogeny. It seems unlikely, for example, that every possible combination of L and H chains yields a functional antibody molecule because in vitro L and H reassociation experiments show that certain hybrid molecules (formed from L and H chains derived from different antibodies) are relatively unstable. Similarly, association of V and J (or V, D, and J) is conceivably not completely random. Evidence of striking bias in the selection of VH genes in fetal pre-B hybridomas has been mentioned. In mice these hybridomas are biased toward the use of genes from the VH7183 and VQ52 families. In addition, fetal and newborn VDJ junctions show a paucity of N nucleotides and a tendency to form VDJ junctions across short stretches of sequence identity between the recombining sequences (homology-mediated recombination, discussed earlier in this chapter). Both of these effects reduce diversity at the recombination junctions, perhaps reflecting a mechanism that ensures the production of certain antibodies advantageous for young individuals. The neonatal bias toward usage of VH7183 and VQ52 families is not observed in adult B cells, but this bias raises the possibility that other less striking recombination biases may exist in adults, reducing the actual diversity below that calculated on simplistic assumptions. It has been reported, for instance, that mouse Jκ rearrangements use Jκ1 and Jκ2 preferentially (398,399), whereas human B cells use JH4 preferentially (400) so that the combinatorial contribution of the available J regions to diversity is probably less than it would be if all were used equally frequently.

Somatic Mutation

Some early arguments suggesting the existence of somatic mutation in antibody genes were based on claims that estimates of combinatorial diversity (as computed along the lines of the above exercise) were, although vast, still too small to account for the observed number of different antibodies. The latter number might be estimated from the percentage of B cells binding a particular antigen and the number of different antibodies—within that binding specificity—that could be distinguished by isoelectric focusing, idiotype characteristics, or analysis of the fine specificity of antigen binding. Such arguments based on global evaluations of diversity were superseded by studies of systems with restricted diversity, in which germline and expressed repertoire can be more reliably compared; these studies generated convincing evidence for somatic mutation. The brief account below summarizes some of the major features of somatic mutation; this topic is discussed in detail in Chapter 24.

Early Evidence for Somatic Mutation

Analyses of amino acid sequences of murine λ1 chains from myeloma antibodies provided the first strong support for somatic mutation, even before the era of recombinant DNA analysis. Thus, when the amino acid sequences of λ1 chains produced by 21 independently derived myelomas were analyzed (401,402), 12 were found to be identical, representing a prototype Vλ1 sequence. The remaining variants were each unique, generally differing from the prototype sequence by single amino acid substitutions that could be accounted for by single base changes. Significantly, all but one of the amino acid substitutions were unique to a single variant sequence. The investigators concluded that the prototype sequence corresponded to a single germline gene, whereas the variants arose by somatic mutation of this gene. This interpretation seemed consistent with the observation that each variant sequence occurred only once, whereas several occurrences of the same sequence might have been expected if there were several germline Vλ1 genes. Now that gene cloning has confirmed that there is only a single Vλ1 gene, the identification of the variants as products of somatic mutation has been verified.

Subsequent studies led to similar conclusions for mouse Vκ or VH systems involving small V families whose germline members could be readily cloned. An example of such a system is the relatively restricted murine antibody response to phosphorylcholine (PC). Sequence analysis of a panel of PC-binding hybridomas and myelomas expressing a similar VH sequence showed that all IgM antibodies shared a single prototype sequence (403). In contrast, some IgA and most IgG VH regions showed scattered amino acid substitutions with respect to the prototype sequence. All of the sequence variants were unique to single cell lines. By analogy to the Vλ system discussed above, these comparisons suggested that the prototype sequences reflected a germline gene, whereas the variants were products of diverse somatic mutations. A search of the four germline VH-region genes homologous to the prototype expressed VH gene showed only one gene that could have served as a precursor for the PC-binding VH regions; and this one matched the prototype sequence exactly (404). The fact that the variant VH sequences were seen only in IgA and IgG, not in IgM, is consistent with the fact that IgM is characteristically produced early in the immune response, whereas somatic mutation occurs later in the response overlapping the stage of isotype switching; other studies have shown that somatic mutation can be seen in IgM at a low frequency.

Role of Hypermutation in Immune Responses

To understand the role of somatic mutation in the antibody response, several groups have studied the extent of somatic mutation at different times after the immunization of mice. Studies of the responses to p-azophenylarsonate (Ars), phosphorylcholine, influenza hemagglutinin, oxazalone, and several other antigens have all indicated that the initial response after primary immunization is contributed by antibodies showing no somatic mutation. About 1 week after immunization, mutated sequences begin to be observed, increasing during the next week or so. Booster immunizations yield sequences showing additional mutations.

Many hybridomas made late in the immune response produce mutated antibodies with a higher antigen affinity than the unmutated (sometimes loosely called germline) antibodies made early after immunization. The shift to higher affinity is a phenomenon long recognized at the level of (polyclonal) antisera and has been termed "affinity maturation." This phenomenon can now be explained as the result of an evolutionary mechanism selecting antibodies of progressively higher affinity from the pool of randomly mutated V sequences. According to this model, at the time of initial antigen exposure an animal has a set of B lymphocytes expressing germline (unmutated) versions of Ig sequences resulting from gene rearrangements that occurred before immunization. Because of the diversity of available VH, D, JH, VL, and JL sequences as well as the impressive recombinational potential described earlier, some B cells will express Ig molecules capable of binding the antigen with modest affinity. These cells are stimulated (by antigen binding) to

proliferate and to secrete antibody. Activated B cells located in lymphoid follicles also bind antigen and receive T-cell help; at some point in the response the somatic hypermutation machinery is activated in these cells, generating random mutations in the Ig genes of stimulated cells in the GCs. Many of these mutations can be expected to reduce the resulting antibody's affinity for antigen; indeed, such mutated antibodies with markedly reduced affinity have been demonstrated (405), as have mutated antibodies that have acquired autoantibody specificity (406). As antigen clearance reduces antigen concentrations seen by the lymphocytes, only the cells displaying high affinity antibody will be effectively stimulated by antigen; cells displaying lower affinity antibodies or antibodies with affinity for self antigens may be subjected to programmed cell death (apoptosis) (407–409). The preferential proliferation of the high-affinity cells and their maturation to secreting plasma cells will be reflected in an increase in the average affinity of the antibodies in the serum. These high-affinity cells will be left as the predominant population to be represented as memory cells when antigen exposure ceases; they thus can induce the rapid, high-affinity response on secondary antigen exposure. In this model the driving force for affinity maturation—analogous to natural selection in the evolution of species—is selection for high antibody affinity in the face of low antigen concentration. The importance of this selective force is suggested by the observation that repeated injection of antigen can inhibit affinity maturation, as though by abolishing the selective pressure for high affinity (410).

Cellular Context of Somatic Mutation

Somatic mutations occur primarily in B cells of the GCs of lymphoid tissues (411,412), particularly in a subpopulation of B cells known as centroblasts. These cells proliferate in the "dark zone" of the GC and bear characteristic surface markers, including IgD, CD38, and the receptor for peanut agglutinin (413,414). Each GC appears to be populated by a small number of antigen-specific founder B cells (412) and an unusual Thy-1–negative T-cell population, also antigen specific (415). The GC environment promotes contact between the B cell and follicular dendritic cells (FDCs) which store, process, and present antigen, and T-lymphocytes, which activate somatic mutation in part via CD40–CD40L interaction (416). Proliferating GC centroblasts give rise to centrocytes, which are programmed for apoptosis unless they are rescued by FDC-presented antigen and T-cell activation via CD40 engagement (409,417). It is at this stage where positive selection for high-affinity antibodies occurs via apoptosis of cells expressing low-affinity antibodies, yet paradoxically apoptosis is also promoted by soluble antigen, perhaps functioning to select against autoantibodies (408,418,419). As mentioned earlier, receptor editing may be another fate for autoantibody-producing cells in GCs. The features of antigen signaling that select for survival versus apoptosis or editing are not fully understood. Susceptibility of GC cell populations to apoptosis is correlated with their expression of Fas, Bax, p53, and c-*myc*, all of which promote apoptosis, as well as down-regulation of the apoptosis suppressor Bcl-2. B cells of mice with engineered up-regulation of Bcl-2 expression can escape selection against autoreactivity (419a).

Germinal center B cells may undergo several successive cycles of mutation followed by selection. Such a scheme is suggested by the sequence analysis of mutated Ig genes PCR-amplified from single cells microdissected from a histologic section of a GC (420); resulting sequences can be organized into genealogical trees consistent with several stages of somatic mutation. Additional evidence for successive mutations has been reported in purified memory B cells (421). A computer simulation has affirmed the high efficiency of alternating periods of somatic mutation and mutation-free selection as a strategy for generating high-affinity antibodies (422). Despite the evidence that somatic mutation occurs normally in GCs, mice lacking histologically detectable GCs as a result of lymphotoxin-α-deficiency are capable of affinity maturation and somatic hypermutation (422a).

Distribution and Targeting of Mutations

To explore the mechanism of somatic hypermutation, several groups have examined the distribution of mutations around Ig genes by comparing the sequences of somatically mutated rearranged genes to their germline precursors. Mutations occur not only in sequence derived from the germline V coding sequence, but also in the J region and nearby flanking intron sequence derived from upstream of the C-region gene. The somatic mutations seem to cluster in the V(D)J region, extending upstream no further than the RNA initiation site (with few exceptions) and tapering off downstream to define a target domain of about 1.5 kb. Therefore, for VDJ units involving the 3′ JH4 segment, mutations extend farther downstream than for units involving JH1 (423,424). The focal nature of the mutations suggests a specific Ig hypermutation mechanism that recognizes some feature of the DNA in or near the VDJ sequence as a target for mutations.

Exactly what feature of the V(D)J locus targets the hypermutation machinery is not understood. Unrearranged Vκ, VH, and D regions are generally not mutated, suggesting that the functional target probably includes elements contributed by both V and J (425–427); however, unrearranged Vλ regions can be mutated (428). This difference may be related to the fact that unrearranged Vλ genes are transcribed in B cells (429), whereas unrearranged Vκ genes are not (275). Therefore, the element that is contributed by V(D)J recombination in support of Vκ and VH hypermutation may be the proximity of the V-region promoters to enhancers lying near the C region, which can increase transcription. The specific chromosomal location of Ig genes does not seem to be necessary for hypermutation because transgenic mice carrying a rearranged expressible Ig gene—presumably inserted randomly in the genome—show somatic mutations in copies of the transgene cloned from hybridomas (430).

The appearance of hypermutation in transgenes has allowed further experimentation on the sequence requirements for mutation through studies of the effects of altered transgene structure on the mutation rate. The importance of transcription in targeting hypermutation is reinforced by studies of transgenic constructs engineered with or without either of the two transcriptional enhancers associated with the κ locus: the "intronic" enhancer lying between the Jκ segments and Cκ, and the downstream enhancer lying 3′ from Cκ. Rearranged κ transgenes that included the downstream enhancer and other downstream elements were more highly transcribed and better somatic mutation targets than similar constructs lacking these regions (431,432,432a,432b), whereas removal of the intronic enhancer essentially abolished hypermutation (432). Furthermore, a VκJκ-Cκ transgene in which a duplicate copy of the Vκ promoter was engineered upstream from the Cκ region was found to incur mutations over 1.5-kb domains downstream from both promoters; the extra promoter created a new mutation domain extending into the Cκ region (433). However, the promotion

hypermutation does not seem to be specific to Ig promoters because replacement of the Vκ promoter with the β-globin promoter did not abolish hypermutation (432); non-Ig enhancers also can promote hypermutation (434). Furthermore, the Vκ coding sequence can be replaced by a human β-globin gene or prokaryotic *neo* or *gpt* gene without affecting the hypermutation rate downstream from the promoter (435). In contrast, a similar transgenic construct in which the Vκ gene was replaced by the CD72 gene was not targeted for hypermutation despite high levels of transcription (436), and even a highly expressed Vλ-Cλ transgene was not mutated (437). To summarize, it appears that transcription is necessary but not sufficient for targeting hypermutation, and additional requirements have not been defined as of this writing. Currently available data leave open the possibility that targeting of V genes for somatic mutation is not very specific and that some non-Ig genes that are transcribed in GC B cells also may be subject to mutation (438). Somatic mutations observed in the *bcl6* gene may represent an example of this phenomenon (438a).

Because mutations are not confined to hypervariable (CDR) regions and sometimes even occur in introns, it is apparent that the hypermutation mechanism does not distinguish coding from noncoding regions, let alone hypervariable regions from framework. The apparent clustering of mutations in the CDRs of sequenced Igs may be partly a result of selection for cells expressing primarily CDR mutations, either because framework alterations interfere with the basic folding of the protein or because CDR mutations can lead to higher affinity for antigen and thus stronger activation to clonal expansion, as discussed above. However, in Ig genes that are not selected for function (e.g., nonproductively rearranged VDJ alleles or passenger transgenes engineered with stop codons to prevent expression as a protein), mutational hot spots as well as cold spots have been recognized, apparently due to local DNA features that may promote or suppress somatic mutation within the domain of DNA targeted for hypermutation. It is possible that evolution has selected for sequences that create mutational hot spots in CDR regions to enhance the potential for diversity generation in the parts of the protein critical for antigen contact (439,440).

Molecular Mechanism of Hypermutation

The molecular mechanism of the mutations remains obscure. The observed mutations have shown little about what may have caused them. All four nucleotides have been targets for mutation, and all have been products. Both transitions (purine–purine and pyrimidine–pyrimidine interchanges) and transversions (purine–pyrimidine interchanges) have been observed, with apparent preferential targeting of G-C base pairs (441). Small insertions and deletions rarely occur. Significantly, in an unselected passenger Vκ transgene, A and G nucleotides were mutated more frequently on the coding strand than on the noncoding strand (442); this strand polarity—also observed in human VH regions (443)—suggests that the mutation mechanism may be affected by a process that can distinguish between the strands, such as transcription through the V region.

One report has argued that somatic mutations in an expressed mouse VH gene occurred by gene conversion, i.e., the clustered changes were templated by a nearby related VH region whose sequence agrees with all the observed mutations in the expressed gene (444). Apparent gene conversion also was observed in a mouse strain carrying a transgenic gene construct designed to optimize the possibility of conversion events (445). Clearly gene conversion seems to play a major role in somatic diversification of chicken and rabbit V genes (446,447) and probably pig V genes as well (448), and it may play a role in the evolutionary diversification of the germline repertoire (449,450); but no further evidence supporting a role for gene conversion in somatic diversification of murine or human Ig genes has been reported, even in cases where such conversion events might be easily detected (451). Indeed, gene conversion could not explain many examples of somatic mutation—e.g., in Vλ1 genes and in the J regions and associated introns—because no closely related but different sequences are present in germline DNA that could donate the mutated nucleotides found in these regions of rearranged genes. This argument also applies to the prokaryotic transgenes targeted for hypermutation in the experiments described above.

The observation that the sequences near somatically mutated nucleotides in Ig genes include direct repeats and palindromic sequences has led to the suggestion that these may play a role in somatic mutation (452). It also has been proposed that mutations may be generated by an error-prone polymerase during repair of nicks or gaps in the DNA (453). Because patients or mice with several defects in DNA repair seem competent for Ig somatic hypermutation, the affected genes are apparently dispensable for this process (454,454a).

A recent model envisions the mutations as a consequence of transcription-coupled repair (433). In one version of this model, a mutator protein specific to GC B cells might load onto the transcriptional complex at the promoter and cause pausing of the complex at various positions; this pausing would induce gratuitous transcription-coupled repair that would occasionally produce errors. Multiple rounds of transcription in each cell could offer repeated opportunities for mutation by this mechanism. In each round of transcription, the mutator protein would fall off the transcription complex as a stochastic event during progression of the complex downstream, thus accounting for the irregular decline in the mutation frequency at increasing distances from the promoter. Such a model would be consistent with the strong correlation between the transcription initiation site and the 5′ boundary of mutations (454b,454c).

The product of the mismatch repair gene *Pms2* (homologous to mutL in *E. coli*) has been implicated in somatic mutation by a recent experiment: a *Pms2* knockout allele was bred into a mouse strain—the quasimonoclonal or QM mouse—engineered with rearranged VκJκ and VDJ genes knocked-in to the respective germline loci by homologous recombination (454d). Although the *Pms2* mutation causes a general increase in mutation rate in most tissues (454e), the immunoglobulin genes in B cells showed significantly fewer somatic mutations than were seen in the QM mouse with normal *Pms2* (454d), suggesting that *Pms2* activity contributes to Ig gene hypermutation. Mismatch repair machinery could hypothetically participate in Ig gene hypermutation by switching its usual preference for correcting the newly synthesized strand, instead preserving any mutations in this strand by "correcting" the opposite strand.

An important, but as yet unclarified, role for IgD in somatic mutation is suggested by the observation that mice with a homozygous targeted disruption of their Cδ gene were impaired, although not completely deficient, in affinity maturation (455). Conversely, an $IgM^{-}IgD^{+}$ subset of GC B cells from human tonsils were found to accumulate extremely high numbers of somatic mutations (456).

Investigations of somatic mutation should be facilitated by recently described systems for observing the process in vitro in primary B cells (454f,457) or cell lines (441,458) and by the development of rapid methods for detecting somatic mutation (459).

Immunoglobulin Gene Evolution: Varying Roles for V Gene Assembly

Evolution of the Immunoglobulin Superfamily and V Assembly Recombination

The three families of Ig genes (κ, λ, and H chains) and the closely related four families of TCR genes (α, β, γ, and δ) clearly provide a powerful and flexible molecular defense mechanism that is valuable for survival in the face of the diverse pathogenic microorganisms that abound in our environment. How did such a complex and elegant system evolve? One obvious approach to elucidating Ig gene evolution is to infer genetic history from comparisons of the Ig gene systems in different modern species. The more ancient history of these genes may be approached by examining homologous non-Ig genes. The ever-growing number of non-Ig genes that demonstrate sequence similarity, and therefore presumed homology, to the Ig genes has become known as the Ig superfamily (460,461). This family name reflects the fact that the Ig genes were the first members to be sequenced and does not imply a functional relationship of the superfamily to Igs or to the immune system. The hallmark of the Ig superfamily is the Ig domain: approximately 100 amino acids, generally encoded in a single exon, and including an internal disulfide loop spanning roughly 60 to 70 amino acids. Despite some rather tenuous primary sequence similarities, the Ig domains are all assumed to share approximately the same three-dimensional structure found in Igs, comprising seven roughly parallel strands forming two layers of β-pleated sheets. This assumption has been confirmed for several members of the superfamily, including β2-microglobulin, CD4, TCR-α and -β chains, and the α3 domain of MHC class I.

Almost all of the Ig superfamily members are surface proteins that function by contacting other surface proteins in cell–cell interactions. Because Ig superfamily members mediating such interactions are found in even the most primitive metazoan organisms—e.g., cell adhesion molecules in slime molds (462)—the Ig domain is likely to be truly ancient, significantly predating the function of superfamily members in defense against microbial invasion. On the other hand, several invertebrate superfamily members have been described that do play a role in microbial defense, e.g., the molluscan defense molecule (463) and insect hemolin (464), which may even share with Ig genes some features of gene regulation by Rel family transcription factors (465). Such examples suggest that some members of the superfamily may have functioned in a primitive immune system in a common ancestor of molluscs, insects, and vertebrates.

Although the Ig domains of most currently known superfamily members are encoded in single exons, several examples (e.g., CD4, N-CAM, and the Xenopus CTX protein) are encoded in two separate exons (466,467). It is uncertain whether this structure reflects an origin of the Ig domain from association of two primordial half-domains by exon shuffling (with later loss of the intron in most current superfamily genes) or the introduction of an intron into a preexisting single-exon Ig domain. Several of the distantly related superfamily members appear to be more C-like or V-like, suggesting that they arose after the divergence between the primordial C domain and V domain. However, this sequence of events is not definitively established, and examples of both C- and V-like genes are known that show the divided half-domain structure.

Do the separated V-region elements (V, D, and J) found in modern Ig (and TCR) genes reflect an association of originally unrelated sequences or fragmentation of elements that were contiguous in an ancestral gene? Suggestive observations bearing on this question have come from an analysis of the CD8 gene. A genealogic relationship between the CD8 antigen and κ Ig is suggested not only by sequence similarity but also by the linkage of their genes on chromosome 6 of mice and chromosome 2 of humans. The CD8 gene has been found to include a segment of J-like sequence contiguous with the V-like sequence (468), suggesting that V and J sequences may have been contiguous in a primordial ancestor gene. The presently observed separation of V and J may then have resulted from insertion of DNA between them by a transposition-like event (Fig. 17), as originally proposed by Sakano et al. (6). A second similar insertion may have separated D sequence from the germline V, as shown in the Fig. 17, although an alternative hypothesis involving a single insertion event also has been proposed (469). In order for the V, (D), and J regions to be reassembled after they were rendered noncontiguous by the transposition event, one would need to assume the prior or simultaneous development of a mechanism—presumably based on the RAG proteins—for V(D)J recombination. Because there is no evidence for RAG-like genes in primitive species without V(D)J recombination, simultaneous acquisition of the RAG genes and the insertion separating germline V, (D), and J elements appears reasonable, lending favor to the speculation that the RAG genes may have been carried on a transposonlike element—flanked by 7-mer and 9-mer RSS repeats or both ends—that inserted into a primitive V region. Such a transposon might have derived by lateral transfer from a prokaryotic element. Presumably the sequence inserted roughly 400 million years ago into the genome of a primitive cartilaginous fish because RAG genes (208), as well as V(D)J recombination of Ig and TCR genes are found in modern sharks and all higher vertebrates examined (470–472); but none of these parameters are found in the slightly more primitive lamprey and hagfish. Interestingly, the shark RAG 1 gene shows sequence similarity to the INT (integrase) gene of phage λ (as well as to the yeast DNA repair proteins RAD 16 and 18 and the human breast cancer susceptibility gene BRCA1) whereas RAG-2 shows sequence similarity to the bacterial integration host factor gene. The similarities to modern prokaryotic genes with a recombination-related function strengthens the hypothesis of a prokaryotic source for the RAG genes. Moreover, as noted earlier, the mechanism of RAG-catalyzed DNA rearrangement—with a hairpin intermediate—bears some similarity to prokaryotic DNA recombination mechanisms.

V(D)J recombination may have provided primordial Ig superfamily genes with their first potential for somatic diversification, i.e., variable junctions resulting from the flexibility of the recombination position. Presumably, as soon as diversity became functionally important for recognition of specific foreign antigen, mechanisms for clonal activation would have been developed and allelic exclusion would have become important to focus the specificity of the response. Conceivably these features arose before the divergence between Ig and TCR genes.

Diverse Evolutionary Mechanisms for Diversity

Ig genes of the shark, chicken, and rabbit provide interesting contrasts to the more familiar evolutionary paths taken by mice and humans. The shark H-chain locus consists of multiple duplicates of ~10-kb units containing separated V, D, J, and CH elements (65,473). The V, D, and J elements are associated with recombination signal elements similar to their mammalian homologs. Sequence comparisons between duplicated units demonstrate differences not only between the germline V genes but between t

CH genes as well. If VDJ recombination occurs only within one of these repeat units, as has been presumed, then diversity would derive from junctional flexibility but not from combinatorial multiplication; sharks also would lack the level of diversity afforded by the isotype switch because a particular V region would always be associated with a specific CH region. This limited system allows the shark to mount specific antibody responses, but these responses show remarkably little variation between individuals, although somatic mutation does occur. Presumably, the mechanism for clonal selection is operative in sharks.

Diversity generation in chickens follows a still different scheme. The λ system has been particularly well studied (474,475) and in the germline consists of a Vλ1 gene 1.7 kb upstream from a typical Jλ-Cλ unit. All the expressed λ protein apparently derives from VJ recombination involving this Vλ1 gene. Upstream of Vλ1 lie 25 Vλ pseudogenes. These pseudogenes cannot themselves encode viable V regions and do not rearrange with the Jλ segment. However, they contribute to diversity by donating stretches of their sequence to the rearranged Vλ1 by a somatic gene conversion process; expressed Vλ sequences show multiple patches of sequence that differ from Vλ1

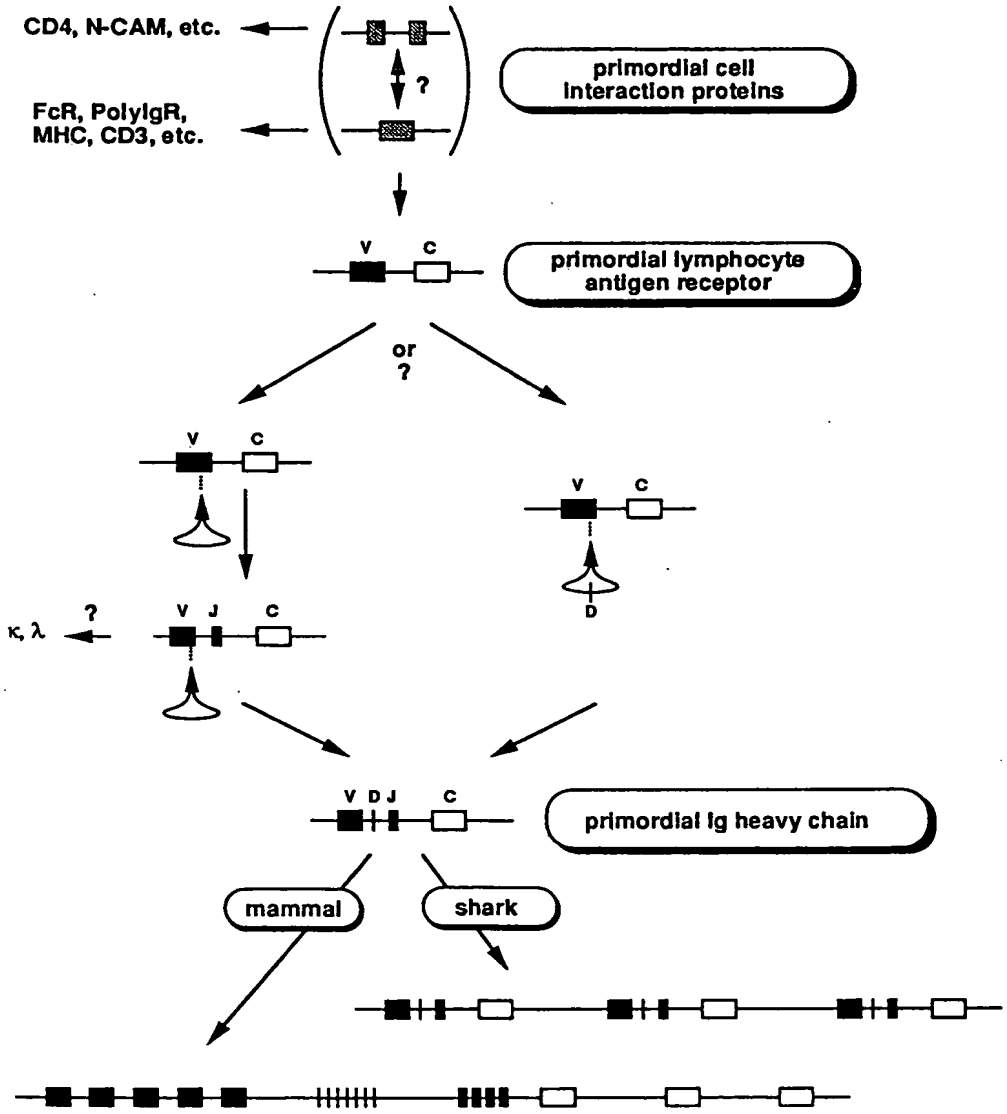


FIG. 17. Evolution of Ig genes. The Ig superfamily presumably evolved from a primordial cell interaction domain by multiple rounds of duplication followed by individual mutation and specialization by different duplicated copies. Because both Ig and TCR systems share specialized V and C domains and V assembly recombination, these features probably evolved before divergence of these two antigen receptor systems. The evolution of the split V region (requiring assembly of V, D, and J to form a functional domain) could have resulted from a single event inserting a D sequence and flanking DNA between V and J of a primordial V domain (pathway on right) or by two separate insertion events (pathway on left). The primordial H-chain gene evolved by different pathways of duplication in shark and mammalian lineages. Sharks show duplication of the entire VDJC unit, whereas in mammals separate duplication of each of these elements occurred.

but precisely match specific pseudogene sequences. Thus, in contrast to the somatic hypermutation observed in mice and humans, individual expressed chicken λ genes show no evidence of random point mutations, and no sequence alterations are found in the intron sequences flanking the rearranged VJ unit. Although combinatorial joining diversity is completely absent in this system, the chicken is capable of a highly heterogeneous λ response as a result of multiple rounds of gene conversion events operating in different regions of the V segment. A similar gene conversion mechanism is important in generating diversity in the chicken H-chain system.

Rabbit Igs might have been expected to follow the schema demonstrated for the homologous loci from the two most intensively studied mammals (mice and humans), but the facts are more interesting. A particularly puzzling feature of rabbit Ig relates to expression of VH allotypes. From 70% to 90% of rabbit antibodies display one of three serologically defined allotypes known as a1, a2, or a3. Rabbits that express one predominant allotype pass this characteristic to their progeny as though a single gene with three alleles were being transmitted as a Mendelian codominant trait; however, Southern blots of rabbit genomic DNA showed several hundred VH-hybridizing bands. How could the simple inheritance of allotype expression be explained given the large number of VH genes? The answer, as demonstrated primarily by Knight and colleagues, is that rearrangements of the most D-proximal VH region, designated VH1, account for most of the H chains expressed in rabbits, and this VH gene encodes the specific amino acids that define the VHa allotype (476). The other VH-region segments contribute to diversity primarily by gene conversion events that alter the VH1 sequence (447); somatic point mutations apparently occur as well (476a). These upstream VH regions may occasionally rearrange productively, perhaps accounting for the 10% to 30% of VHa allotype-negative antibodies in normal rabbits. The potential for such recombination is suggested by a strain of rabbits (Alicia) in which the VH1 gene was deleted; this strain nevertheless makes normal amounts of antibody, most of which is VHa allotype negative. Gene conversion also contributes to diversity in bovine Igs (477).

Undoubtedly, examination of the Ig genes of other organisms will provide additional details of the evolution of these remarkable loci and a better understanding of the differing strategies for the generation of diversity. Such studies also should help to elucidate the evolutionary relationship between Ig genes, the homologous TCR genes, and other Ig superfamily members not involved in immune defense.

REGULATION OF IMMUNOGLOBULIN GENE EXPRESSION

General Principles of Gene Regulation

The mechanisms that regulate the expression of Ig genes have been under intense investigation in recent years as part of a wide effort to understand development and differentiation in molecular terms. Immunoglobulins are synthesized only by B-lymphocytes; even within this lineage these proteins are made in differing amounts at different developmental stages. Although rates of protein synthesis can be regulated at the levels of mRNA transcription, processing, transport, stability, and translation, most attention has focused on transcription because this seems to be the limiting step in most systems that have been examined [although changes in mRNA stability can clearly play an important role (478)]. The gene loci encoding lymphocyte antigen receptors (Ig and TCR) are unique in that the complete genes do not exist in the earliest stages of lymphocyte maturation; only the germline precursors are present. Thus, the regulation of Ig gene expression must be integrated with the progress of Ig gene rearrangements. These processes are further intertwined because as discussed elsewhere in this chapter, transcription is apparently required for Ig gene rearrangement, both for V(D)J assembly and isotype switch recombination. Thus, the transcriptional regulation of nonrearranged loci also merits analysis.

Cis Regulation

Gene transcription can be regulated by cis influences—dependent on the DNA sequence of genetic elements attached to a gene—and trans influences, dependent on the environment of the gene. For most genes, regulatory studies have initially focused on the cis elements that regulate gene expression. Some insights have been gained by examining how gene expression is affected by spontaneous mutations or deletions of regulatory sequences in cells or animals. However, most advances have been made by inserting putative regulatory elements into DNA constructs containing a reporter gene—one whose expression can be conveniently assayed—and then transfecting the constructs back into eukaryotic cells; the function of the putative regulatory elements is then tested by assaying for reporter gene expression. In some experiments the assays are performed only 2 to 3 days after transfection, an interval so short that most of the DNA remains in an unstable episomal form; these are known as transient transfections. In contrast, other experiments are designed to produce stable transfectants in which the engineered DNA construct becomes inserted into the cell chromosomes. As an alternative to transfection of cells, similar constructs can be introduced into the mouse genome, thereby creating strains of transgenic mice. The expression of the introduced transgene can then be assessed in a variety of tissues in the animal to examine whether the candidate regulatory element can program the same pattern of tissue-specific expression as that observed for the gene from which the element was derived.

Through such transfection and transgene experiments three major classes of eukaryotic cis regulatory elements have been defined. A promoter is a DNA segment that is located near the transcriptional initiation site and that promotes the initiation of RNA transcription in a specific direction, i.e., toward the coding sequence of the gene. An enhancer is a DNA segment that can stimulate transcription when positioned at variable distances from the transcription initiation site and in either orientation. A silencer downregulates transcription, operating (like an enhancer) in both orientations and over variable distances via mechanisms not thoroughly understood. All three kinds of elements are generally active in only certain cell types and thus participate in regulating the tissue-specific expression of the associated gene. Two other types of cis elements have been characterized in eukaryotic chromosomes and should be noted. Matrix attachment regions (MARs) attach DNA to the chromosomal scaffold proteins and may promote local unpairing of the DNA strands (479,480). Locus control regions (LCRs), first discovered in the β-globin locus (481), are complex regulatory regions that are composed of smaller elements that individually have enhancer function. LCRs affect chromatin structure and gene activity over longer distances than enhancers are thought to act. Operationally they are defined by their ability—when tested in transgenic constructs—to program associated reporter genes for expression independent of the position of integration into chromosomal DNA; in contrast, constructs without LCRs generally are expressed at widely different levels in different transgenic mouse strains depending on integration site.

Figure 18 provides an overview of the currently known regulatory sequences of the Ig loci in the mouse (similar regions have been reported for most of the homologous human loci). Promoters are present in the flanking DNA just upstream of each V gene in all three loci: κ, λ, and H chain. In plasmacytomas only the promoter of the rearranged V region is active, whereas similar promoters of unrearranged upstream Vκ or VH regions are inactive. This observation provoked a search for an additional regulatory sequence downstream from the J that might activate the promoter of the adjacent rearranged V region. The J-C regions of the κ and H-chain loci were screened for regulatory regions, and enhancers were found in J-C introns of both loci. (The J-C introns of λ loci apparently lack enhancers.) Near the intronic enhancers of the κ and IgH loci, silencer regions have been reported that may inhibit the activity of the associated enhancers in non-B cells. After the discovery of intron enhancers, two observations led to expectations of additional enhancers 3′ from the C-region genes. First, several myelomas were found to have undergone spontaneous deletions that eliminated the J-C intronic enhancer of the expressed H-chain gene, but the myelomas nonetheless continued to express this gene at normal levels; these observations were consistent with the presence of an additional enhancer in the DNA that had not been deleted. Second, enhancers were discovered downstream from C-region genes in the related family of TCR genes. Subsequent investigation uncovered enhancers 3′ from κ and λ C-region genes, as well as an enhancer 3′ from the Cα gene, the most downstream constant gene in the H-chain locus.

These enhancers, silencers, and the V-region promoters may be sufficient to explain the transcription of complete, assembled Ig genes; but additional germline or sterile transcripts are transcribed from Ig C-region genes that are being activated for V assembly or isotype switch rearrangements. These transcripts are also controlled by promoters (Fig. 18), which in some cases have been found to be critical for regulation of the associated DNA rearrangement.

Promoters, enhancers, and silencers are composed of clusters of several short sequence motifs, each of which can be recognized by a specific nuclear protein (or proteins). Some of these motifs are present in more than one enhancer or may even be shared between enhancers and promoters. In the discussion below, several of the important murine regulatory regions and their functional motifs are described, along with nuclear protein families known to regulate Ig gene expression by binding to these motifs. Each murine regulatory region has been found to have an apparent homolog in humans, often with many of the same nuclear binding motifs conserved. The presence of multiple motifs in a given enhancer complicates the analysis of the role of any one motif. Engineered mutations in a particular motif often have very little effect on the activity of the complete enhancer, and sometimes an artificial construct containing a single functional motif often shows no enhancer activity on its own. Two strategies have been used to demonstrate the function of such motifs. In a construct with an enhancer fragment in which all but a few motifs have been deleted, the contribution of each remaining element is often detectable through the effects of mutations. Alternatively, an artificial enhancer containing several multimerized copies of a single motif may have enhancer activity when a single copy does not. The proteins that bind to enhancer motifs mediate the regulatory function by promoting (or inhibiting, in the case of silencers) the assembly of transcriptional machinery at the promoter. The proposed interactions between proteins binding enhancer and promoter imply that the intervening DNA forms a large loop. Many regulatory proteins are present in the nuclei of only certain tissues or cell types, a fact that can in principle explain the cell type specificity of the transcription of particular genes. External stimuli that up- or downregulate Ig gene expression (e.g., cytokines or antigen binding) typically work by altering the amounts or activity of certain DNA-binding proteins.

Types of Trans Effects

Alteration in the nuclear content of DNA binding proteins that interact with cis regulatory elements represents a well-studied mechanism for trans regulation of gene expression, but other approaches to investigating trans regulation should also be mentioned. One correlate of gene activation that formally falls in the class of trans effects is the altered chromatin environment of DNA in expressed genes that is often detectable by nuclease sensitivity experiments. In these experiments isolated nuclei are treated with varying concentrations of DNase I (or a variety of other nucleases, including restriction endonucleases) and the DNA is then purified, digested with a restriction enzyme, and analyzed by Southern blotting using a hybridization probe for the genes under study. In general, when the nuclei are derived from cells expressing a particular gene, that gene is more sensitive to DNase I than unexpressed genes in the same cells; i.e., a Southern blot band carrying the expressed gene can be abolished by treatment with low concentrations of DNase that leave unexpressed genes (or their Southern blot bands) relatively unaffected. In addition, appropriate Southern blot strategies show that certain segments of DNA in expressed genes may be hypersensitive to DNase I; these segments tend to coincide with regulatory regions of genes where sequence-specific binding by regulatory proteins blocks access of these DNA regions to nucleosomes, rendering them accessible to nucleases.

Another chromatin correlate of gene activation is the extent of DNA methylation. Most cytosine residues within CpG dinucleotides are methylated in mammalian DNA, but genes that are actively expressed in a particular cell generally appear relatively undermethylated in that cell type (482). The extent of CpG methylation can conveniently be estimated for that subset of CpG dinucleotides that fall within the sequence CCGG, which is the recognition site for two restriction endonucleases: *Msp* I cuts at this site regardless of the methylation status of the internal cytosines, whereas *Hpa* II cuts only the completely demethylated site. Southern blot strategies that exploit this difference have been used to compare methylation in active and inactive genes. Both μ and κ C-region genes have been shown to be sensitive to DNase I and undermethylated in pre-B cells, B cells, and plasma cells but DNase resistant in nonlymphoid cells (483–487). Cells that are undergoing isotype switching show undermethylation of the C-region genes that are the targets of switch recombination (483,488); such undermethylation correlates with synthesis of germline transcripts from these CH genes. Of V-region genes in B cells, only the rearranged and transcribed V gene generally shows the undermethylation and DNase sensitivity characteristic of active genes (484,489). In the IgH locus, DNase I hypersensitivity sites have been found overlapping the intronic (486) and downstream enhancers (490,491). In the κ gene, hypersensitivity sites occur at the promoter and enhancer as well as at a site 5′ from the enhancer (492,493).

Methods of Studying Trans-Acting Proteins Binding to cis Regulatory Motifs

Recent studies have investigated trans-acting proteins identified by their interaction with known cis-acting promoters or enhancers. In vitro binding of nuclear proteins to specific regulatory sequence elements can be detected by several techniques, some of which can be

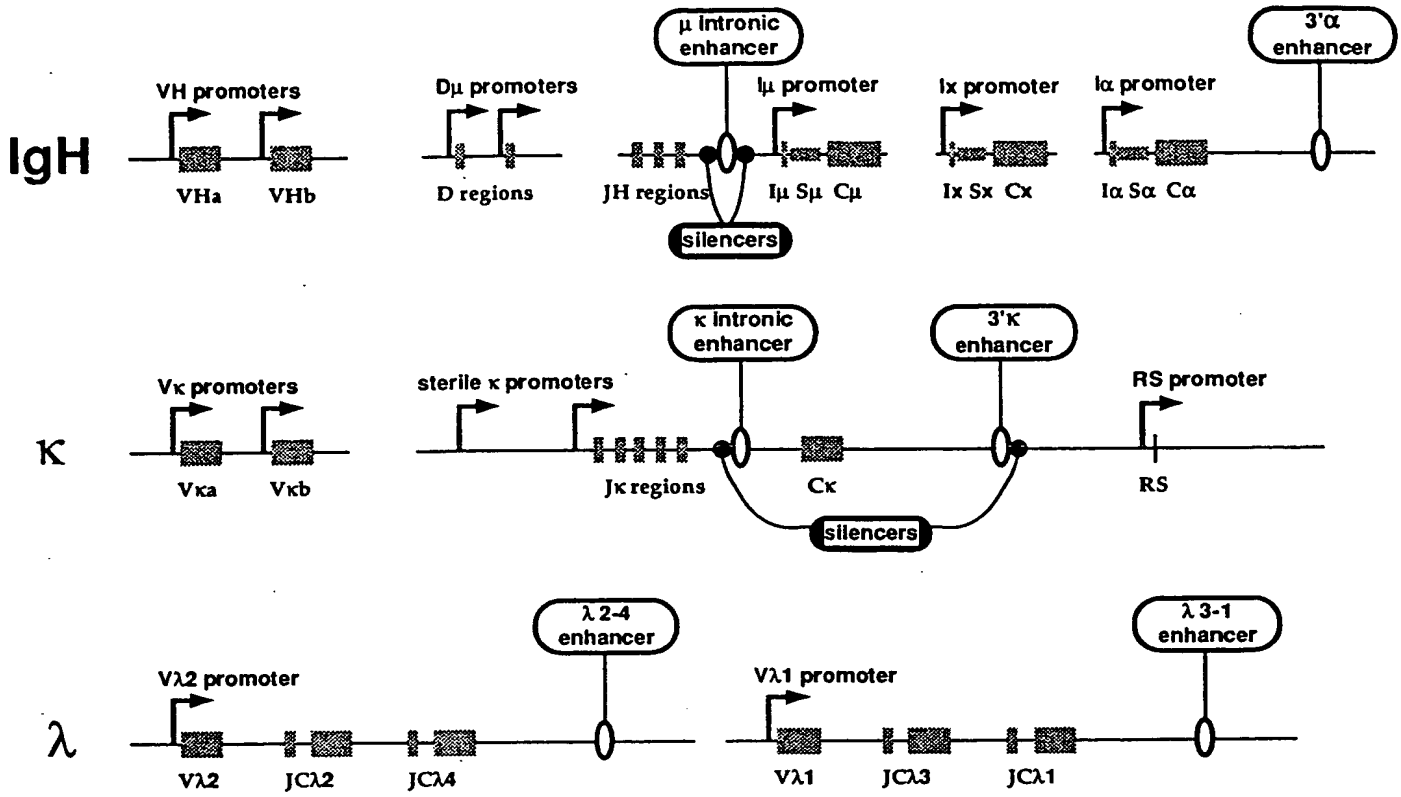


FIG. 18. Enhancers and promoters of the murine Ig loci. Schematic maps (not to scale) of the three murine Ig loci are shown: IgH (*top*), κ (*middle*), and λ (*bottom*). The six known Ig enhancers are shown as vertical ellipses, the four silencer regions by black circles, and the various promoters by arrows indicating the direction of transcription. The enhancer reported upstream from murine DQ52 (572) is not shown in the graphic image.

used to assess sequence-specific binding even in crude protein mixtures. The simplest technique is the EMSA. In this assay a short (typically 30-300 bp) radioactively labeled double-stranded DNA fragment is allowed to interact with a mixture of proteins extracted from cell nuclei by a salt solution; when the DNA is then electrophoresed in an acrylamide gel, binding of protein(s) to the DNA can be detected by the retarded mobility of the protein–DNA complexes in the gel in comparison with the mobility of the free DNA probe. Sequence specificity of the retarded band must be demonstrated by showing (a) that its intensity can be diminished by adding to the incubation mixture an unlabeled competitor oligonucleotide identical to the probe sequence, but (b) that a similar amount of oligonucleotide of unrelated sequence is without effect. Retarded complexes can be identified as containing an already characterized protein if an antibody to that protein specifically eliminates or supershifts (i.e., further retards the electrophoretic mobility) of the band. Another powerful technique, DNA footprinting, allows the visualization of the specific DNA sequence covered by a bound protein. In this technique a protein preparation is allowed to bind to a fragment of DNA that has been radioactively labeled on one end of one strand. The DNA-protein mixture is then treated with DNase under conditions so mild that on average each strand will be nicked by the enzyme only once; then the DNA is purified from the incubated proteins and electrophoresed on a denaturing acrylamide gel (along with size markers) in order to detect the positions of DNase-induced nicks in the radiolabeled strand. A nuclear protein that can bind tightly to the radioactive DNA fragment during the initial DNase incubation step protects the region of the DNA that it contacts from nuclease attack, and so the position of the bound protein can be inferred from a region of the fragment that is relatively free of nicks (the footprint).

Once a protein has been detected that binds to a critical regulatory element in the DNA, detailed study of the protein requires molecular cloning of its gene. Two main strategies have been used for such cloning. In one approach the protein is first purified by classic fractionation procedures. EMSA or DNase footprinting assays can be used to follow the binding protein through fractionation steps. Typically, the purification includes an affinity column in which the DNA sequence representing the binding target is fixed to the column bed; the specifically interacting protein binds this DNA sequence with high affinity, separating it from contaminating material. When the protein is pure, amino acid sequences are obtained from tryptic fragments; these sequences are used to design DNA probes that can be used to isolate clones from a cDNA library. An alternative cloning strategy (494) bypasses the protein purification procedure. From a cell expressing the binding protein, a cDNA library is constructed using the vector λgt11, a bacteriophage engineered to allow transcription and translation of insert cDNA sequences in infected bacteria. A library of viral plaques imprinted onto membrane filters is screened by soaking the filters in a solution containing the target DNA binding sequence as a radioactively labeled fragment. A plaque that expresses a cDNA encoding the binding protein is able to bind the probe and thus creates a radioactive spot on an autoradiograph of the filters. Clones identified by their position on the filters are then isolated for study.

The fact that a purified nuclear protein binds in a sequence-specific manner to a regulatory DNA sequence does not prove that this protein mediates the regulatory function of the DNA sequence. However, a functional role for an enhancer-binding protein can be inferred if transfection of a clone encoding the protein induces transcription of a cotransfected reporter gene linked to the enhancer/promoter motif that is bound by the cloned protein. Such experiments have verified the function of several Ig promoter- and enhancer-binding proteins, which can therefore be considered transcription factors. Some of these occur only in B-lymphocytes and thus can account in part for the B-cell specificity of Ig gene expression. Others are more widespread. Many of these regulators of Ig genes are homologous to mammalian oncogenes as well as to genes of drosophila and yeast, suggesting ancient evolutionary origins and fundamental importance of these proteins in the regulation of cellular metabolism.

Several of these proteins will be discussed below in connection with the cis regulatory elements with which they interact. Most of this discussion is based on analysis of murine Ig genes, which have been examined most extensively.

Cis-Acting Elements in V-Region Promoters

The Octamer Motif

The transcription of assembled Ig genes initiates upstream from the V gene sequences. The promoters that regulate this initiation are, by virtue of their upstream positions, present in each germline V-region gene even before V assembly recombination. Like many eukaryotic genes, most V gene promoters contain a TATA site about 25 bp 5′ from the initiation site; TATA sites serve as binding sites for the transcription factor TFIID and related proteins and thereby play a role in specifying the exact position where RNA transcription initiates. The only other conserved feature of all classes of Ig V promoters (i.e., κ, λ, and H chain) is an octamer ATTTGCAT that is found associated with Vκ and Vλ genes, whereas the inverted complement ATGCAAAT is found 5′ from VH genes (495,496). (The same sequence is sometimes identified as a decamer TNATTTGCAT or the complementary ATGCAAATNA.) Vκ promoters generally include only the octamer plus the TATA box, whereas VH and Vλ promoters can include both of these motifs as well as other characteristic regulatory elements. The conservation of the octamer element in Ig V promoters suggested that it might play an important role in Ig gene function; indeed, when constructs containing this motif were analyzed by transfection, it became clear that the octamer is critical in conferring B-cell specificity to the promoter. Deletions or mutations in the octamer cause dramatically reduced B cell–specific promoter activity when tested either in Ig gene constructs or in heterologous genes transfected into B cells (497–501). The octamer also has been shown to be required for optimal in vitro transcription by B-cell nuclear extracts, whereas it had no effect on transcription by HeLa extracts (502). Octamers appear in the promoters of several B cell–specific genes other than Ig, including B29 (Igβ) (503), CD21 (504), and CD20 (505). A multimerized octamer can act as a B cell–specific enhancer (506).

A puzzling feature of transcriptional regulation by the octamer is that this sequence is also a functional component of promoters of several other genes whose transcription is not B cell specific. These include the herpes thymidine kinase gene, histone H2B genes, and U1 and U2 small nuclear RNA genes. To understand the puzzling relationship between octamers in Ig and non-Ig promoters, several laboratories have undertaken analyses of the trans-acting nuclear proteins that bind to these elements. Two such proteins—designated Oct-1 and Oct-2—have been extensively characterized, initially by EMSA experiments and subsequently by gene cloning (507–510). These two proteins show differing tissue distributions. Most cells make Oct-1 (also known as OTF-1 and NFA-1), but only B cells and a few other cell types (notably, activated T cells) make Oct-2 (OTF-2, NFA-2). Several additional octamer-binding proteins specific to other tissues (e.g., in neural cells or embryonic stem cells) have been reported. The Oct proteins share a similar 160–amino acid DNA binding domain, which explains their virtually identical binding specificity. Amino acid sequences similar to this DNA binding domain have been found in several other nuclear proteins that bind to motifs resembling the octamer. This binding domain thus defines a family of nuclear factors, which has been designated the POU family (pronounced "pow"), named for the three factors in which this conserved domain was first noted: *P*it-1, *O*ct-1/2, and the nematode gene *u*nc86. The domain includes a 75-to 80-residue POU-specific domain (POU_S), a short flexible linker, and a 60–amino acid segment (POU_H) homologous to the homeobox domain. (Homeoboxes were first recognized in genes regulating drosophila development, but more recently have been noted in genes throughout the animal kingdom and even in plants.) The POU_S domain contacts the ATGC part of the ATGCAAAT sequence, whereas the POU_H domain contacts the AAAT segment (511).

The Oct proteins have been demonstrated to be transcription factors by experiments in which the corresponding genes were cotransfected into fibroblasts or HeLa cells along with reporter gene constructs driven by octamer-containing promoters. Critical activation regions, necessary for the Oct proteins to stimulate transcription, have been deduced from the effects of deletions and mutations placed in different regions of Oct proteins (512) and the effects of swapping (through genetic engineering) various domains between Oct-1, Oct-2, and other POU proteins (513–515). On the N-terminal side of the POU domain, Oct-1 and Oct-2 both contain a glutamine-rich activation region, but on the C-terminal side Oct-2 contains a feature missing from Oct-1: an activation region rich in serines, threonines, and prolines. Apparently the C-terminal differences are functionally important because swapping the N-terminal domains between Oct-1 and Oct-2 has little effect, whereas replacing the C-terminal domain of Oct-1 by that of Oct-2 confers a distinctive property of Oct-2: the ability to activate transcription from multiple octamer motifs functioning as either a promoter or an enhancer (506).

The B-cell specificity of Oct-2 suggested that this factor might be important for activity of the octamer motif in V promoters, and some evidence supports this inference. However, targeted disruption of the Oct-2 genes in a B-cell line (516) produced little effect on the expression of either endogenous Ig genes or a transfected gene driven by an octamer-containing promoter. Furthermore, although homozygous Oct-2 knockout mice (517) die without obvious pathology within a few hours of birth, they contain roughly normal numbers of B cells, which respond to activation by a T-cell clone with near normal cell proliferation and Ig secretion (518). These results suggest that Oct-2 is unnecessary for the V-region promoter activity required either for early B-cell development or for T cell–activated Ig secretion, possibly because of the redundant role of Oct-1 for these processes. On the other hand, cultured B cells from the homozygous Oct-2 knockout animals showed marked defects in LPS-plus-cytokine–activated Ig secretion and in anti-IgM–induced proliferation, suggesting a role for Oct-2–dependent proteins in these signaling pathways.

The B-cell specificity of Oct proteins is complicated by their interactions with additional proteins. One such protein—designated

octamer coactivator from B cells (OCA-B)—was originally detected in affinity-purified preparations of either Oct-1 or Oct-2 as a factor necessary for optimal in vitro transcription from a Vκ promoter (519). The purified protein has now been cloned by several laboratories, which use several different names for it: OCA-B (520), OBF-1 (521), and Bob-1 (522). Binding of OCA-B to the octamer/Oct–protein complex apparently stimulates transcription through a transcriptional activation domain of OCA-B (523). The OCA-B protein binds to the POU domain of either Oct-1 or Oct-2, but also may contact the DNA at the fifth base of the ATGCAAAT sequence; oligonucleotides with alterations at that position bind the Oct proteins normally but cannot form a complex with OCA-B, nor can a reporter construct mutated at that position show OCA-B–induced stimulation of transcription (523,524). An important role of this protein for in vivo regulation of Ig production is suggested by the effects of OCA-B disruption by gene targeting (525,526). OCA-B knockout mice seem healthy and are fertile but show defects in B-cell maturation and Ig production. The number of mature B cells in the spleen is reduced, and the response to immunization is dramatically impaired, with reduced proliferation and a severe decrease in IgG, IgA, and IgE, apparently due to decreased Ig gene transcription in B cells that have undergone isotype switch recombination. GCs are not formed in these mice. Some of these effects are apparently mediated by decreased Ig gene transcription, whereas others may result from interference with OCA-B–dependent expression of other genes. Purified OCA-B seems to bind preferentially to Oct-1 rather than Oct-2 (527); a second coactivator has been postulated to mediate Oct-2–dependent transactivation (528).

What is the critical feature of octamer motifs in Ig promoters that confers B-cell specificity when the same motif in an H2B promoter is active ubiquitously? Although the answer is still not known, one possibility is that apart from the TATA box, octamers in Ig promoters are not associated with other important promoter motifs that might allow ubiquitous expression in ubiquitously expressed genes. Consistent with this view is the observation that the insertion of a CCAAT promoter motif to an otherwise lymphoid-specific promoter renders the promoter active in nonlymphoid cells (529). It is also possible that—as in the case of several other coactivators for Oct-1, including VP16 (530,531) and PTF (532)—sequences outside the octamer play a role in discriminating between Ig and other promoters. OCA-B may mediate some of this discrimination because OCA-B, when added to a HeLa-derived in vitro transcription system or when coexpressed in HeLa cells, could coactivate a construct driven by a Vκ promoter much more effectively than a similar construct with an H2B promoter, even though both promoter sequences supported complex formation with OCA-B (520,521). Candidate motifs that might contribute to B-cell specificity of Vκ promoters include sequences downstream from the transcription start site (532a).

In addition to its role in V promoters, the octamer also appears in the H-chain enhancer, where it can contribute to the B cell specificity of constructs transfected into various cell lines (533), although it did not appear critical for enhancer function in transgenic mice (534). This octamer may activate the enhancer under certain conditions of cell stimulation (535) and clearly plays an important role when this region functions as a promoter driving sterile transcripts of the Cμ gene (536). Additional DNA segments that are similar, but not identical, to the octamer have been found in several other regulatory regions of Ig genes, e.g., upstream of the mouse κ intron enhancer (537). However, the functional importance of most of these octamerlike motifs has not been demonstrated. The mechanism of regulation by the Oct proteins is likely to be considerably more complex than outlined here because of the existence of several isoforms resulting from alternative RNA splicing (538–540), several phosphorylation states critical to protein function (513,541), and the ability of both OCA-B and the Oct proteins to interact with other regulatory factors ((520,542,543).

Other Elements of Ig V Promoters

Although unusual Vκ promoters that lack efficient octame motifs have been reported to attain promoter activity through ar alternative pyrimidine-rich element designated κY (544), and a motif binding early B-cell factor may contribute to some Vκ promoters (545), Vκ promoters are typically composed only of the octamer plus TATA box. In contrast, Vλ and VH promoters routinely contain additional functional elements besides the octame and TATA box, some of which are briefly described below. In VH promoters a heptamer CTCATGA generally lying 2 to 22 bp 5 from the octamer was found to be well conserved and required fo optimum promoter activity (546). Surprisingly, although thi sequence bears little resemblance to the conserved octamer ATG CAAAT, it appears to bind in vitro to both Oct-1 and Oct- (547–551). The heptamer binds these proteins with lower intrinsi affinity but shows cooperative interaction with occupancy of a adjacent octamer site. Cooperativity also can be demonstrated at functional level by in vitro transcription experiments (551) Another element showing some sequence conservation in VH promoters and a role in optimal promoter function is a polypyrimidin tract located between 0 and 46 bp upstream from the heptame (546). A motif that includes a polypyrimidine tract (GGAACCTC CCCC) has been identified as a component required for optima function of the MOPC141 VH promoter (552). This motif, whicl was designated the N element, was found to bind a novel tran scription factor of ubiquitous distribution. The relationshi between the N element and the pyrimidine-rich κY motif is no known at present. A motif (TTANGTAA) conserved in many VH regions binds to C/EBP factors, originally identified as binding t the E motif in the μ enhancer. This motif is required for optima transcription of VH promoter–driven transfected constructs i vivo, and the purified binding protein stimulates transcription fror such promoters in vitro (553). One final VH element deserve mention. In an investigation of the mechanism by which treatmer with the lymphokine IL-5 plus antigen upregulates Ig H-chai mRNA, Webb et al. (554) detected an A/T-rich element betwee 125 and 250 bp 5′ from the VHS107 start site that could mediat increased transcription by these agents. In an EMSA experimer this element produced a band that was upregulated in extract obtained from cells treated with IL-5 plus antigen. In its A/T rich ness, the element resembles MARs, and a cloned protein corre sponding to this binding activity—designated B-cell regulator c IgH transcription (Bright) (555)—partitions partly with the insolu ble chromatin matrix. The significance of the Bright binding sit for VH function is uncertain because most VH regions lack simila sequences within the 5′ flanking region as far as has bee sequenced; furthermore, a transgenic construct driven by a relate VH promoter deleted for Bright binding was still expressed in lymphocyte-specific manner (556). Additional response elemen may be discovered as the mechanisms of Ig transcription response to other manipulations (including other lymphokines) a investigated. To speculate further, the several VH promoter ele ments that are absent from VL promoters may facilitate the earl transcription of germline VH genes, allowing VDJ recombinatio to occur at a time when VL genes are transcriptionally inactiv

however, much additional evidence would be necessary to support such a hypothesis. It is also possible that variations in the content or spacing of different elements in different VH promoters may differentially regulate V gene transcription, thereby influencing the frequency with which specific V regions are rearranged and utilized in the Ig repertoire (557–559).

Vλ promoters have received less study than Vκ or VH. When the two major murine Vλ genes were studied, the Vλ2 promoter was found to contain octamer and TATA plus an additional functional element located upstream from the octamer and not precisely conserved in Vλ1. This element, CACGTGAC, is identical to that recognized by the protein USF (upstream stimulatory factor) (560). USF is a ubiquitous transcription factor originally isolated based on its role in regulating the major late promoter of adenovirus, but since then found to regulate a wide variety of cellular genes. It belongs to the family of helix-loop-helix (HLH) transcription factors that is described in more detail below. When Vλ2 promoter activity was studied in an in vitro system with a B-cell nuclear extract, transcription was found to be reduced by passage of the extract through an anti-USF antibody column but could be restored by the addition of purified USF (561). Thus, it appears that the murine Vλ2 promoter includes a functional USF motif. Little is currently known about the functional components of human Vλ promoters or of other murine Vλ promoters except that most of them contain octamer-like sequences and TATA boxes.

Promoters of Sterile Transcripts

Ig RNA transcripts that are sterile (i.e., do not encode a complete Ig protein) are produced before V(D)J and isotype switch rearrangements, as discussed earlier in this chapter. The sterile transcription, by altering the accessibility of the DNA, may constitute part of the regulation of the corresponding gene rearrangements. Sterile transcripts including unrearranged V regions (274,562,563), Cκ (305,564,565), or Cμ (566,567) regions were initially detected, and more recently sterile transcripts of most of the downstream CH genes of the murine and human loci have been characterized. Salient features of the regulation of these transcripts are briefly outlined below.

Sterile V-Region Transcripts

The promoter of a particular V region is assumed to be identical whether sterile or mature—i.e., V(D)JC—transcripts are produced. Therefore, the earlier discussion of V-region promoters probably applies to sterile transcripts. It remains to be determined, however, what mechanisms allow the promiscuous transcription of sterile V transcripts during the developmental stage when V(D)J rearrangements are occurring, but later shut off the promoters of all but the rearranged V regions (274,275).

Sterile Cμ Transcripts

Two types of sterile μ transcripts have been described. In the first type, transcription can initiate at heterogeneous positions near the 5′ end of the JH-Cμ intronic enhancer (Eμ). When the resulting RNA transcripts are spliced to the Cμ1 exon, an Iμ exon (intron-derived) remains attached to the RNA encoding Cμ. The promoter of Iμ-Cμ transcripts has been found to be coincident with the Eμ enhancer; however, as mentioned above, the octamer motif plays a much more prominent role for this region as a promoter than it does as an enhancer (536). This promoter lacks a TATA box in both murine (536) and human (568) loci. Because the TATA box is generally responsible for establishing a precise transcription start site, its absence from the Iμ promoter probably accounts for the heterogeneity of the 5′ ends of these transcripts. The Iμ exon is remarkable for containing multiple stop codons in all three reading frames. This feature has been hypothesized (567) to protect against the possibility that the Iμ-Cμ transcript might be translated into a protein that could prematurely provide the signal to terminate V(D)J recombination, a signal that is normally generated by a complete H-chain protein (see the discussion on allelic exclusion earlier in this chapter). After switch recombination has produced a composite switch junction (e.g., Sμ-Sγ), the Iμ promoter is retained and remains active, leading to hybrid transcripts such as Iμ-Cγ (569).

The second type of sterile Cμ transcript is derived from loci that have undergone DJ rearrangement (273); these transcripts have the structure DJ-Cμ after splicing out the JH-Cμ intron. These transcripts initiate from promoters that lie upstream of the germline DH elements (570) but have not been fully characterized. For the most JH-proximal murine D region, DQ52, an upstream promoter and enhancer have been reported (571,572). In mice, the subset of DJ-Cμ transcripts in reading frame 2 (RF2) encodes the Dμ protein that suppresses RF2 in expressed H-chain V regions, as described earlier.

Sterile Cκ Transcripts

Two types of sterile Cκ transcripts also have been described: an 8.4-kb primary transcript (564,573) that initiates about 3.5 kb upstream from Jκ1 (and is processed to a 1.1-kb RNA) and a 4.7-kb primary transcript that initiates just upstream from Jκ1 (and is processed to 0.8 kb) (573). Both of these transcripts are found in pre-B cells and are upregulated by exposure to LPS. The 5′ flanking sequences of both initiation sites contain octamer-like sequences (7/8 match to consensus) which are capable of sequence-specific binding to nuclear proteins, binding that can be competed by cold consensus octamer oligonucleotides (573). Close to the initiation site near Jκ1 there are also two binding sites for an additional protein designated KLP, which is B cell specific (574). Further experiments will be necessary to completely delineate the functional elements of these promoters and to understand how they might respond to the presence of a μ protein in the pre-B cell to activate κ transcription and subsequent κ rearrangement. A sterile JCκ transcript has been reported to encode a "V-less" κ protein analogous to the Dμ protein, which might associate with μ H chain on the surface of pre-B cells as an alternative to the VpreB-λ5 SLC complex (575). The function that such a protein might have is unclear; certainly the striking phenotype of λ5 knockout mice indicates that a JCκ protein is not redundant with the SLC.

One final transcript from the κ locus should be mentioned for completeness. To test the generality of the principle that gene rearrangements are associated with transcription of the recombination targets, the RS element involved in deletion of the Cκ gene (discussed earlier in this chapter) was examined for transcriptional activity. A transcript derived from this region was found in some pre-B cells (576). The regulation of expression of this transcript remains to be explored.

Sterile I-CH Transcripts

The regulation of sterile transcripts of downstream CH regions is being actively investigated because this regulation may be important for understanding the mechanism by which numerous factors influence the selection of specific isotypes expressed in

particular immune responses (as described earlier, and discussed more fully in Chapter 23). In this view the promoter for the sterile transcript of a given isotype may be expected to contain a unique combination of motifs mediating the action of antigen, various cytokines, and other T-cell influences that are known to regulate switching to that isotype. The regulation of the IgE response has been studied intensively because of its clinical implications for allergy, and it may represent an illustrative example.

The switching of B cells from the production of μ H chains to ε is highly dependent on the cytokine IL-4, as demonstrated by the abolition of IgE synthesis in IL-4 knockout mice (577). In experiments in vitro, switching of splenic B cells to ε production requires IL-4 in the presence of an additional signal that can be supplied by several mitogenic treatments, including LPS, anti-CD40, T-cell membranes, etc. In vitro production of ε H-chain protein is preceded by synthesis of sterile Iε-Cε RNA transcripts—also known as germline ε (Gε) transcripts—which initiate at multiple start sites, apparently owing to the absence of a TATA site in the promoter (578). Significantly, the Iε promoter can confer IL-4 inducibility to reporter gene constructs. The minimum sequence with this capacity contains binding sites for two known nuclear proteins (579). One of these is STAT6, a member of the family of signal transducers and activators of transcription, which transduce signals from many cytokine receptors to mediate transcriptional regulation. STAT6 is activated by engagement of the IL-4 receptor and is required for the IL-4 effect on Iε transcription and switching to Cε, as shown by experiments in cells lacking STAT6 and in STAT6 knockout mice (580–582). The other component required in the minimal IL-4 responsive element of the Iε promoter is a binding site for the CAAT/enhancer binding protein (C/EBP) family of transcription factors. This family includes C/EBPα, expressed constitutively in liver cells, and NF-IL6 (C/EBPβ), which mediates the action of LPS and inflammatory cytokines such as IL-1, tumor necrosis factor-α (TNF-α), and IL-6. Another member of this family is the widely expressed C/EBPγ (also known as Ig/EBP), which lacks a transcriptional activator domain but can act as a transdominant negative inhibitor of other C/EBP family members by heterodimerizing with them (583). Changing ratios of different members of this family in B-cell development contribute to regulated expression of VH promoters, and the intronic enhancers of the κ and IgH loci (584).

In addition to the STAT6 and C/EBP binding sites, two nearby motifs closer to the Iε initiation sites contribute to optimal IL-4 inducibility of the promoter and also mediate the synergistic response of the promoter to CD40 engagement (585); these sites bind to the complex family of proteins known as NF-κB, which is described below in connection with the intronic κ enhancer. Elimination of both NF-κB sites from the promoter inhibits IL-4 inducibility, consistent with the absence of Iε sterile transcripts and switching to Cε expression in mice with targeted deletion of the NF-κB component p50 (586).

An additional level of IL-4 control of the promoter apparently results from an A/T-rich sequence overlapping with some of the Iε transcriptional initiation sites. This sequence confers repression of the promoter, apparently due to binding of the chromosomal protein HMG-I(Y) (587). IL-4 induces phosphorylation of this protein, perhaps thereby decreasing binding affinity and relieving the transcriptional repression (588). An additional component of the murine Iε promoter that contributes to basal activity but not to IL-4 inducibility is a binding site for the B cell–specific activator protein (BSAP) (585,589), a transcription factor described in more detail below. In the human Iε promoter, a BSAP site apparently enhances both IL-4- and CD40-mediated promoter activity (589a). Apart from the promoter, sterile Iε transcription and isotype switching are also regulated by an enhancer lying downstream from the murine Cα gene, as deduced from abnormalities in mice in which this enhancer was replaced by a neomycin resistance gene in all B cells (96): defects were observed in switching to IgE and several IgG isotypes (but not to IgG1), with corresponding decreases in sterile transcripts of the affected isotypes.

Investigations of the regulation of sterile transcripts of other CH genes suggest promoters of similar complexity. In general, these promoters include motifs that act as response elements for signals known to promote switching to the respective isotypes; additional discussion can be found in Chapter 23.

Cis Elements of Ig Gene Enhancers

As pointed out above, enhancers are regulatory elements that stimulate transcription of nearby genes but, in contrast to promoters, can affect transcriptional initiation thousands of base pairs away and in either orientation. Enhancers have been found upstream and downstream from genes and, as is the case in the Ig genes, in introns. Although their position and orientation independence led to a variety of speculative models to explain how these sequences act to stimulate transcription, the prevailing view at present is that, like promoters, enhancers bind to nuclear proteins that facilitate assembly of a transcription initiation complex. As mentioned already, even enhancers that are positioned thousands of base pairs away from the initiation site in terms of linear distance on the DNA sequence can come close to the promoter simply by forming a large loop of DNA that doubles back on itself. Such loops have been observed by electron microscopy in several model systems.

The discussion that follows focuses on murine enhancers, which have been studied in most detail. The human homologs are briefly described at the end of this section.

Heavy-Chain Intronic Enhancer

E-Boxes and Their Binding Proteins

The enhancer located in the JH-Cμ intron was one of the first cellular (nonviral) enhancers recognized, and it continues to be a target of intense study because of its remarkable complexity (as indicated in Fig. 19A). In both the human and murine loci this enhancer, often designated Eμ, lies about 0.5 kb 3′ from the most downstream JH region and appears to be spread over about 0.3 kb. This segment is 5′ from the μ S region and is thus routinely retained on the expressed gene after isotype switch recombination. A major effort has been made to analyze the mechanism of action of this enhancer by analyzing the component functional motifs and their binding proteins that mediate enhancer activity.

In early work aimed at identifying the positions within the murine Eμ that might serve as binding sites for nuclear proteins, Church, Ephrussi, and colleagues (590,591) used an in vivo version of the footprinting method described above, examining the accessibility of the enhancer region to the methylating agent dimethylsulfate (DMS) in B cells versus nonlymphoid cells. These experiments located four clusters of nucleotides demonstrating B cell–specific alterations in DMS reactivity. These clusters defined a consensus octamer CAGGTGGC that appears not only at these

four positions in the μ intronic enhancer (designated μE1, μE2, μE3, and μE4) but also at a fifth position not apparent on the footprint (μE5) as well as at three positions within the mouse κ enhancer (κE1, κE2, and κE3). These motifs have become known as E boxes or E motifs.

The functional significance of E motifs in the μ enhancer has been tested by transfection of DNA constructs containing the enhancer with one or more of these motifs altered by deletion or by clustered mutations. Most of the E motifs were found to be functionally redundant in that constructs containing mutations in single E motifs or even in several pairs of E motifs still retained substantial enhancer activity (533,592), but a construct with mutations in μE1, μE3, and μE4 showed loss of 98% of the normal enhancer activity. EMSA experiments established that nuclear proteins can bind specifically to the E motifs in vitro. Despite the sequence similarity of these motifs to one another, they do not all bind to the same proteins. Furthermore, although these E motifs were first noted as sites of in vivo alterations in DMS reactivity that were B cell specific, the proteins binding to these motifs were detected by EMSA experiments in nuclear extracts from a surprising variety of nonlymphoid sources.

The identification and characterization of E motif binding proteins showed a complex regulatory mechanism involving evolutionarily ancient components. A central role is played by products of the E2A gene, including two forms—known as E12 and E47—resulting from alternative RNA splicing (593). These proteins, and products of the related E2-2 and HEB genes, bind to the μE2, μE4, and μE5 motifs, as well as the similar κE2 motif from the κ intronic enhancer. All of these proteins are members of the HLH family of transcription factors. This family, now including over 200 proteins from species as diverse as drosophila, yeast, and plants, all share a common consensus binding motif—CANNTG—and the HLH domain: two 13–amino acid α-helices separated by an intervening loop. This structure mediates homo- or heterodimer formation between members of this family. Such dimerization is necessary (but not sufficient) for DNA binding; and the binding specificity and affinity for particular E boxes depends on both members of the dimer pairs. As an example, proteins derived from the E2A gene associate with MyoD (or related muscle factors) to form heterodimers that bind with high affinity to E box–like motifs in promoters of muscle-specific genes (594–596). Apparently E2A-encoded proteins, which are expressed virtually ubiquitously, may participate in tissue-specific gene regulation by binding with HLH partners with narrower tissue distribution. However, the B cell–specific E2A factor (known as BCF-1) is a homodimer of E47 subunits, which for unknown reasons seems to form uniquely in B cells despite the wide tissue distribution of E47 (597,598). Besides the HLH dimerization structure, most HLH proteins (including the E2A products) contain an additional element that is necessary for DNA binding: a segment of basic amino acids adjacent to the HLH on its N-terminal side, hence the designation bHLH for the subclass of proteins that have this basic region. In fact, HLH proteins that lack this segment—e.g., the Id group of HLH proteins (599,600))—apparently serve as physiologic inhibitors of E-motif function by dimerizing with HLH proteins and preventing their DNA binding. An additional component of some HLH proteins is a leucine zipper; this is an α-helical structure with several leucines at seven-residue intervals such that they all project from the same side of the helix. HLH proteins that include leucine zippers (bHLH-zip proteins) can dimerize to each other via hydrophobic interactions between the leucines, but they cannot dimerize to HLH proteins lacking the zipper component (like Id, to which they are thus resistant). The bHLH-zip proteins include the Myc proteins (and their heterodimer partner Max) as well as three proteins present in B-cell nuclei that can bind to μE3 and κE3: USF, TFE3, and TFEB.

With this background on the HLH proteins, several features of their regulatory role in B cells may be considered. The critical importance of E2A proteins for B-cell development is highlighted by the phenotype of mice in which this protein was disrupted by homologous recombination (601,602) Homozygous mutant mice develop to term, but most die within a few days of birth. Strikingly, the mice fail to generate any B-lymphocytes, although the T-cell compartment is grossly normal as are other tissues like muscle in which participation of E2A proteins as heterodimers has been documented. Perhaps products of the related E2-2 gene are able to compensate for lack of E2A in muscle but not in the B-lineage. When transgenes expressing E-12 or E47 transcripts were added to the E2A knockout mice, a synergistic action of these two transcripts in B-cell development was apparent (603). Although impairment of Ig E box–dependent enhancer function may contribute to the E2A knockout phenotype—which includes marked inhibition of Iμ transcription and DJH recombination in fetal liver—this is difficult to establish because of the dramatic reduction in other transcripts important for B-cell development, including RAG-1, mb-1, CD19, and λ5. Transcription of the latter two genes is known to be regulated by the transcription factor BSAP, and expression of the gene for BSAP was also significantly reduced. Thus, impaired interactions between E2A proteins and E-box motifs in Eμ may not contribute significantly to the knockout phenotype. Other evidence for a role of E2A products in regulating Ig genes comes from experiments in which an E47 expression vector was transfected into a T-lymphocyte line (604); this caused a dramatic upregulation of Iμ transcription and DJH rearrangement (although some indirect effects may play a role in this system because expression of Oct-2 and both RAG genes was observed to be increased). E47 overexpression also caused Iμ transcription in a transfected fibroblast line (605). Regulation of E47 may be modulated by phosphorylation in non-B cells, which may reduce the ability of this protein to bind to DNA in the B cell–specific homodimer form (606).

E2-2 and HEB are similar to E2A in structure and in that they are expressed in many different cell types, but their roles in Ig expression have been less well studied. Early in B-lineage development E2-2 is more highly expressed than E2A and probably contributes more to E-box binding (607). Homozygous knockouts for E2-2 or HEB (608) showed unexplained perinatal lethality similar to that observed in E2A knockouts, but only a modest decrease in proB cell numbers; so these genes are clearly are less important for B-cell development than is E2A.

The complexity of the function of E boxes in the μ enhancer is illustrated by investigations on the function of a small fragment of the enhancer containing only μE5, μE2, and μE3 (609,610). A tetramer of this fragment is sufficient to induce enhancer activity in constructs transfected into a B cell. In this enhancer fragment, μE3 mediates a significant part of the enhancer function, and the protein that best mediates this effect is the μE3 binding protein TFE3 (611). Indeed, B cells lacking TFE3 show reduced activation of Ig secretion (612). In contrast to the activity of the tetramerized μE5-μE2-μE3 fragment in B cells, this construct showed no activity in fibroblasts; but a similar tetramer lacking μE5 was active in both B cells and fibroblasts, suggesting that μE5 confers inhibition

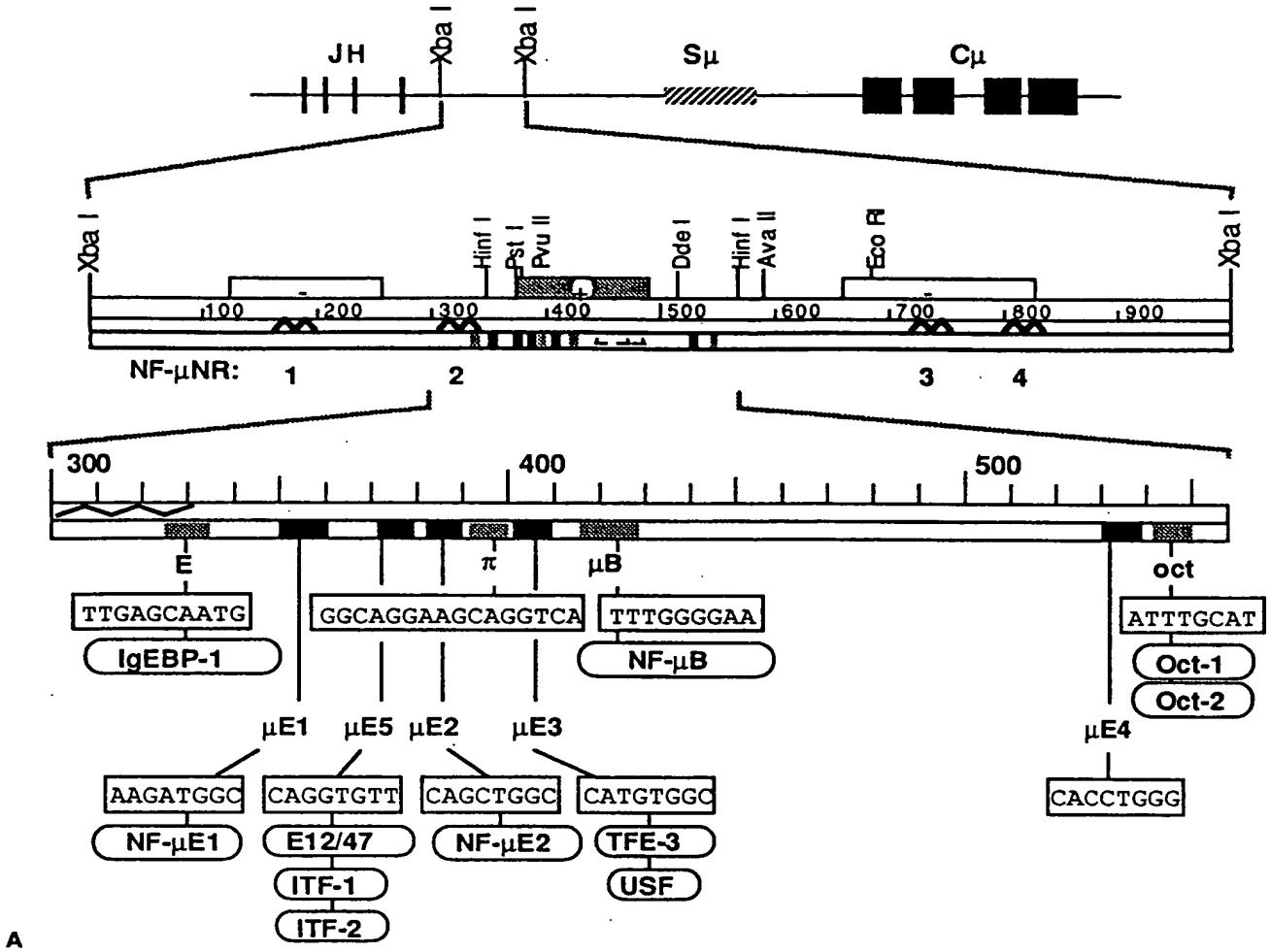


FIG. 19. Four murine Ig gene enhancers. **A:** The H-chain enhancer is located between two *Xba* I sites in the JH-Cμ intron. In the expanded diagram the scale is numbered from the 5′ *Xba* I site; additional restriction sites used by various investigators are shown. The large rectangles above the numbered scale indicate locations of positive or negative regulatory regions identified by Imler et al. (631). The E motifs (*black rectangles*) and other motifs (*shaded rectangle*) are indicated, as are the four binding sites reported for the negative regulatory protein NF-μNR (*sawtooth lines*). The expanded scale at the bottom shows the central motifs with their names, DNA sequence (*rectangles*) and associated binding proteins, where known (*cartouches*).

in non-B cells. This inhibition is apparently mediated by the binding of a non-HLH protein designated ZEB to the μE5 site, allowing inhibition of TFE3. Inhibition by ZEB can be partially reversed by overexpression of the E47-like protein ITF-1, which binds to μE5 and displaces ZEB. Competition between E2A products and ZEB may similarly contribute to the B-cell specificity of the Eμ during B-cell development.

Another mechanism by which Eμ activity mediated by HLH proteins may be regulated is through the Id proteins Id1, Id2, Id3, and Id4. As mentioned above, the Id proteins are dominant negative regulators of bHLH proteins because they can heterodimerize to these transcription factors and prevent them from binding to their cognate motifs in DNA (613). Id1 and Id2 are expressed in pro-B cells at a time when E2A proteins are expressed but are not detectably bound to E boxes. Later these Id proteins are downregulated, apparently allowing bHLH activation of their target regulatory regions (599,600). The model that Id expression can suppress bHLH activation is supported by the phenotype of mice with an Id1 transgene that was designed for late B-cell expression using an mb-1 promoter and the Eμ enhancer (614); these mice showed a marked impairment in B-cell development very similar to the E2A knockout mice described earlier.

A clinically important aspect of the E2A proteins is the capacity of their genes to participate in oncogenic transformation as a consequence of translocations that fuse parts of these genes with foreign genetic material from a different chromosome. The human chromosomal locus of the E2A gene on 19p13 is the site of at least two classes of translocation events in acute lymphocytic leukemia—t(1;19)(q23;p13) and t(17;19)(q22;p13)—that produce oncogenic E2A fusion genes (615,616).

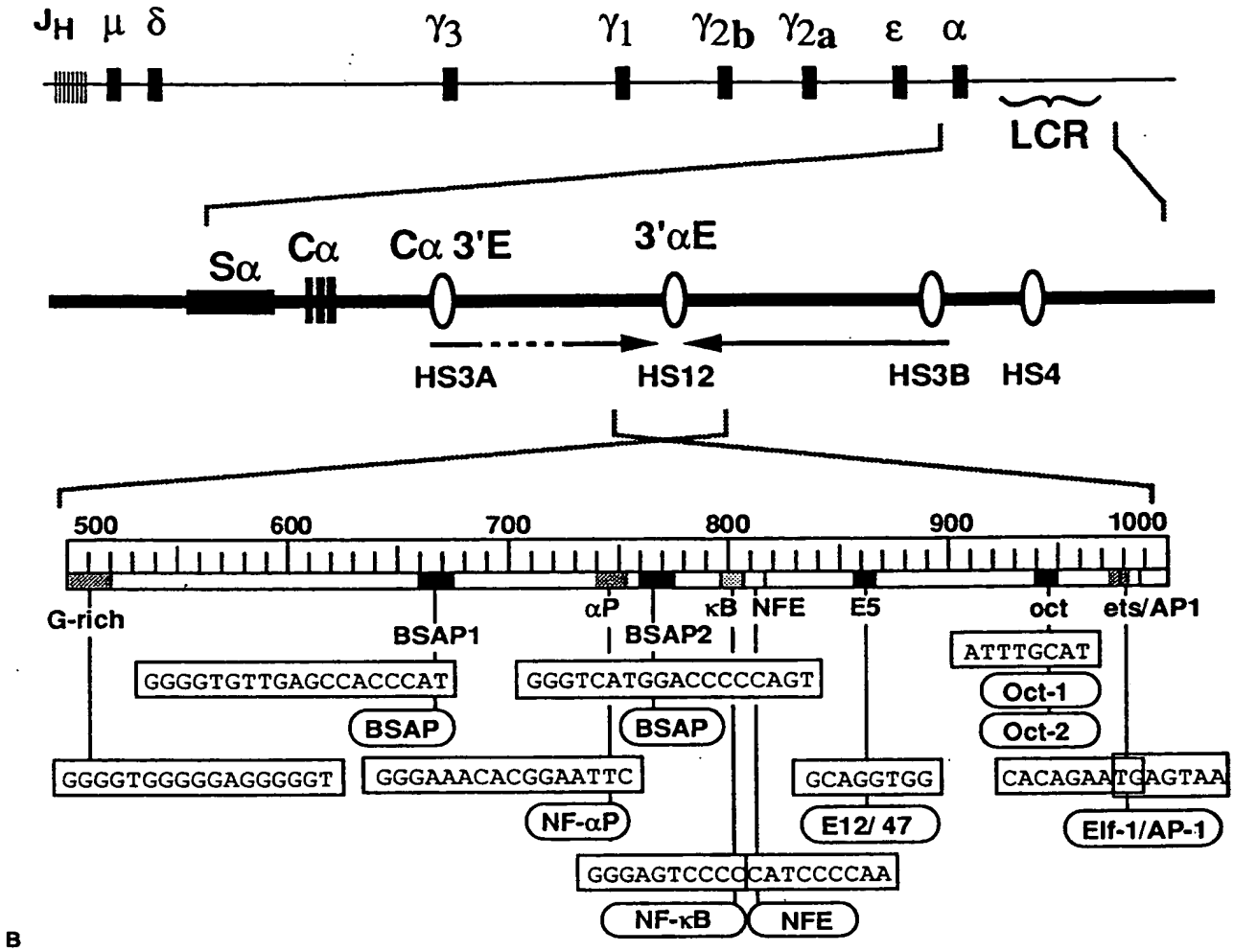
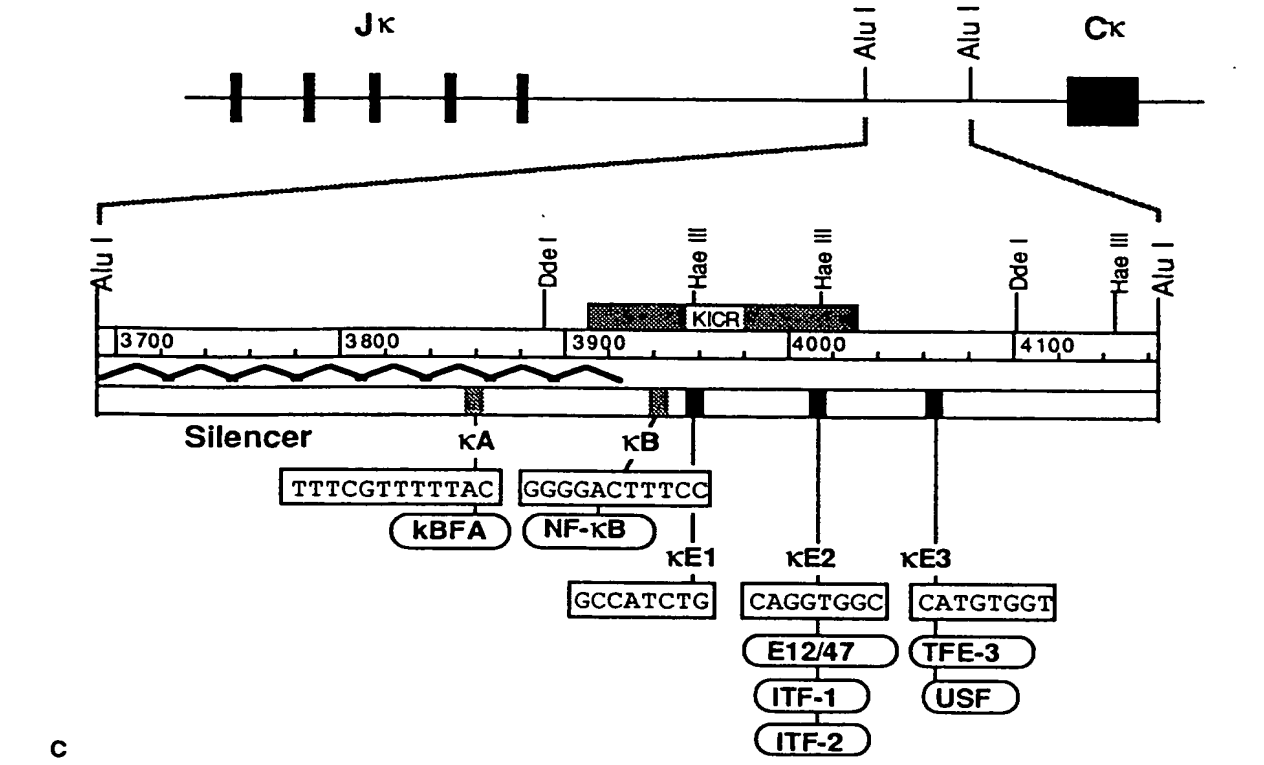


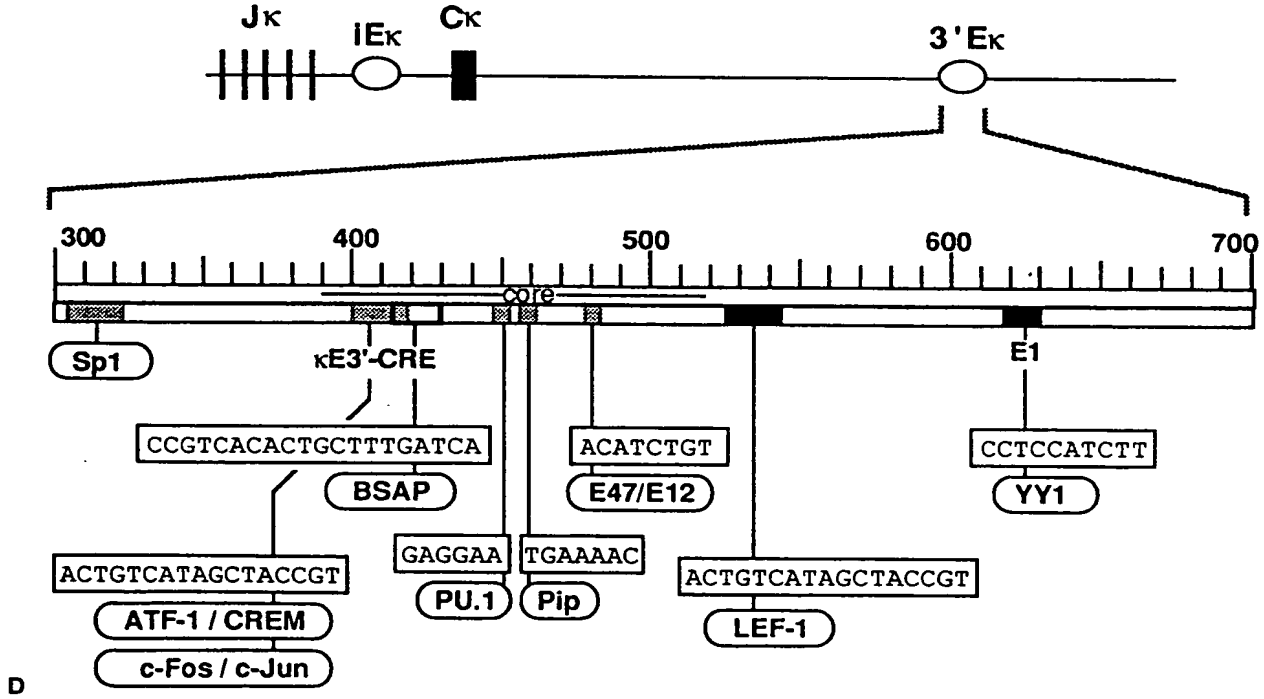
FIG. 19. *Continued.* **B:** Regulatory regions 3′ from murine Cα. The top line shows the location of an LCR downstream from Cα; this is expanded on the next line to display the reported LCR components: HS3A (Ca3′E), HS12 (3′αE), HS3B, and HS4. The arrows show the palindromic structure that surrounds HS12, the major enhancer in the locus. The component's motifs within the HS12 enhancer are shown at the bottom, using the sequence numbering from Dariavach et al. (648). This is the numbering scheme used by most investigators, but it should be noted that this sequence is in the 3′ to 5′ orientation; therefore, the orientation of the third line in the graphic is opposite from that above, and all the boxed sequences are from the antisense strand.

ETS Family Members and Their Role in Eμ

Members of another large family of transcription factors participate in Eμ regulation: the ETS proteins, which bind as monomers to the sites in Eμ known as μA (also called π) and μB (617) (Fig. 19A), as well as to similar sites in the other Ig enhancers. ETS family members share a conserved 85–amino acid DNA binding domain that recognizes DNA motifs generally containing a core GGAA sequence. Mutations at the μA site in transfected reporter constructs suggest that its integrity is crucial for Eμ activity in pre-B cells but not later in B-cell development (618). Several widely expressed ETS proteins, including Ets-1, Ets-2, Erp, and NERF, can bind to μA, although some evidence favors a physiologic role for Elf-1 at this site (619). In contrast, the μB site binds primarily to PU.1, which is expressed only in B cells and macrophages (620). PU.1 appears to be critical for regulation of all three Ig loci and several non-Ig genes, including the mb-1 and J-chain genes. The importance of this protein for lymphoid development has been documented by PU.1 knockout mice; these mice die before birth, showing profound defects in lymphoid and myeloid lineages, although not in erythroid or megakaryocyte cells (621). Although neither μA, μB, nor the intervening μE3 show substantial enhancer activity by themselves when multimerized, the fragment containing these three adjacent motifs does show enhancer activity in B cells (617), and the spacing between the μA and μB elements is critical for activity (622). These findings as well as systematic studies of the in vivo and in vitro interactions between the proteins binding μA, μE3, and μB elements suggest that activity of this minimal enhancer depends on complex mutual interactions between these proteins and the Eμ DNA (623,624). This minimal enhancer shows



C



D

FIG. 19. *Continued.* **C:** The intronic κ enhancer. The murine κ enhancer is located between two *Alu* I sites in the Jκ-Cκ intron. In the expanded diagram the scale is numbered according to a published sequence of this region (824). The E motifs (*black rectangles*) and the κA and κB motifs (*shaded rectangles*) are shown. The shaded rectangle above the scale shows the position of the KICR (671), an island of sequence conservation observed by comparisons between the human, mouse, and rabbit Jκ-Cκ introns. **D:** The 3′ κ enhancer. The motifs and corresponding binding proteins are shown, with the numbering scheme based on Muller et al. (140).

activity in macrophages as well as B cells, but inhibitory influences from flanking E boxes inactivate the enhancer in macrophages, fine tuning the cell type specificity of the enhancer. Similarly complex interactions have been reported between the μE2, μA and μE3 motifs and their binding proteins (624a).

Other Motifs in Eμ

A site in the μ enhancer known as E (unrelated to E boxes) also has been shown to be necessary for optimal enhancer activity (624b,624c). This site binds to members of the C/EBP family (discussed above), which in B cells is represented by Ig/EBP (C/EBPγ) and NF-IL6, with the latter increasing as B-cell development progresses (584). IL-6 is known to upregulate Ig secretion in B-cell lines (625), and the E site is a likely (although not yet documented) participant in this regulation through activation of NF-IL6.

Despite the initial designation of μE1 as an E box, it lacks key criterion now implied by this term, namely the canonical CANNTG motif that seems necessary for HLH protein binding. Instead, the μE1 site is apparently bound by a protein known as YY1 (626) (YinYang1), so called because it can mediate either positive or negative effects on gene expression, depending on circumstances. This protein also binds to a similar site in the 3′ κ enhancer and participates in the regulation of a wide variety of genes in many tissues. YY1 has four zinc fingers (which are zinc-chelating domains found on a subset of DNA binding proteins) and both positive and negative regulatory domains (627,628). The function of this protein in Eμ has not been fully elucidated; mutations of the site decreased enhancer activity in plasmacytoma cells (533) but had no effect in other conditions (629).

Matrix Attachment Regions Flanking Eμ

Several factors apparently contribute to the B-cell specificity of the μ enhancer. As discussed above, two motifs in the enhancer bind to B cell–specific nuclear factors (μB and the octamer), and one motif (μE5) can inhibit function of μ enhancer fragment constructs in fibroblasts. However, several reports have suggested that an additional measure of B-cell specificity is conferred by sequences flanking the enhancer that suppress the activity of the central motifs in nonlymphoid cells (630–632). These suppressive sequences overlap with A/T-rich sequences that flank the core enhancer and have the properties of MARs (633). As discussed earlier, MARs are believed to represent the sites where DNA is tethered to the insoluble protein scaffold of the nuclear matrix, and they have been found near several enhancers (including the κ intron enhancer) and associated with one VH gene. A nuclear protein designated NF-μNR (nuclear factor-μ negative regulator), detected in several cell lines not expressing Ig, has been shown to bind to four A/T-rich μNR elements that lie with one pair in each of the MARs flanking the central enhancer region (Fig. 19A). In transient transfection assays, deletion of these μNR elements from enhancer constructs had little effect on transcription in B cells, which do not express NF-μNR; but in macrophages and T cells, which do express NF-μNR and cannot support activity of the intact enhancer, deletion of the μNR elements activated the enhancer, apparently releasing it from inhibition mediated by NF-μNR (634). This suggested that binding of the MARs to the matrix might be necessary for optimal enhancer function and that NF-μNR might inhibit this interaction in non-B cells. In support of this model, a matrix protein designated MAR-BP1 that might mediate this interaction has been purified from urea-solubilized matrix and has been shown capable of binding to the Eμ-associated MARs; in accordance with the model, this interaction was inhibited by purified NF-μNR (635).

Although the MARs flanking Eμ have been found to contribute little to enhancer activity in constructs transfected into B cells, they are apparently important for Eμ-driven transcription in transgenic constructs (636). When flanked by its MARs, the Eμ demonstrates a defining property of an LCR: it confers position-independent transcription on transgenic constructs integrated at various positions in the genome. Other MARs have been reported to demonstrate this property (637), but the exact relationship between MARs and LCRs is not yet understood. One possibility is that MARs may act by relieving superhelical strain because they correlate with sequences capable of becoming unpaired and nucleating unwinding (479).

There is good evidence that enhancers can stimulate transcription by approximating their binding proteins to promoter-binding proteins, looping out the intervening DNA and forming a three-dimensional transcription factor complex that facilitates formation of a transcription initiation complex. To assess whether enhancers can mediate changes in chromatin structure apart from this promoter–enhancer interaction mechanism, B-cell nuclei from mouse strains harboring a variety of transgenic Eμ constructs without linked eukaryotic promoters were tested for access to DNase and prokaryotic T3 or T7 polymerases (638,639); the constructs contained promoters for the same polymerases, and some contained an MAR from the Eμ flanking region. A minimal Eμ enhancer was found to mediate local factor accessibility, but a MAR was required to extend the accessibility to a promoter 1 kb away from Eμ, implying that MARs can collaborate with an enhancer to generate a domain of chromatin accessibility even without specific interactions between enhancer- and promoter-bound proteins.

Additional insights on MAR function have been gained by studying several other proteins—besides NF-μNR—that can bind to MARs, including SATB1 (640), nucleolin (641) and Bright, as mentioned above (555). But considerable additional work will be required for a comprehensive understanding of how these elements function.

The 3′ α Enhancer and LCR

Discovery of the Enhancer Complex

A complex regulatory locus has been reported to lie downstream from the murine Ig H-chain Cα gene. The existence of a regulatory region in this location was originally inferred when it was found that plasmacytomas that had undergone spontaneous deletions of Eμ nevertheless remained capable of high-level Ig secretion (642–645). Conversely, a myeloma subclone that retained the intronic enhancer but lost a segment of DNA downstream from the murine Cα gene was found to have markedly reduced H-chain gene expression (646). A systematic search in the homologous region of the rat H-chain locus showed an enhancer (647), and a homologous mouse enhancer designated 3′αE was found soon thereafter (648,649) positioned about 16 kb downstream from Cα. The mouse and rat 3′αE segments lie in opposite orientations and are flanked by inverted repeats (648). In addition to the 3′αE, Matthias and Baltimore also reported a weak enhancer in mouse lying only 4 kb downstream from Cα (650).

More recently, Madisen and Groudine (490) analyzed B cell–specific DNase I hypersensitivity downstream from Cα and detected four hypersensitive sites: HS1 and HS2 overlap the previously described 3′αE, whereas HS3 and HS4 lie further down-

stream and identify two new regions with somewhat weaker enhancer activity in transient transfection assays. The HS3 sequence is almost identical to that of the enhancer described by Matthias and Baltimore but in inverted orientation. This reflects the fact that the sequence surrounding the HS12-3′αE is present in the mouse in a long inverted repeat (651,652) (Fig. 19B). When constructs containing HS3, HS12, and HS4 linked to a reporter gene were transfected into a B-cell line, subsequently isolated stable transfectants were found to express the reporter gene in a position-independent manner. This suggested that the three enhancer sequences (HS12, HS3, and HS4) acted together as an LCR.

Component Motifs of the Enhancer Complex

Analyses of the regulatory regions downstream from murine Cα have identified several motifs that bind specific transcription factors to mediate different aspects of regulation of enhancer function. The 3′αE has been found to activate transcription strongly in plasmacytomas, but only weakly in earlier B-lymphoid cells. Part of this developmental change is attributable to a motif known as E5, which matches the E-box consensus binding site—CANNTG—characteristic for members of the bHLH family of transcription factors. The contribution of the E5 site to enhancer activity is inhibited in early stages of development by the dominant negative nuclear regulator Id3, which is expressed in early B-lineage cells but downregulated in plasma cells (653). At least four other motifs in the 3′αE have been reported to contribute to enhancer activity in plasmacytomas but not in early B cells. One site known as αP binds to a member of the ETS family of transcription factors designated NF-αP (654). Another is the octamer motif (ATGCAAAT) common to Ig V-region promoters and several Ig enhancers (655). A third is a κB-like site that binds to members of the NF-κB/Rel family of transcription factors described below (653). The fourth is a G-rich sequence whose function has been demonstrated by mutational analysis but for which binding proteins have not been identified (656). Activity of all four of these sites appears to be regulated by the product of the Pax-5 gene known as BSAP, mentioned earlier. This protein, which binds to two motifs in the 3′αE, is present in early B-lineage cells, in which it suppresses enhancer activity; but its loss in plasmacytomas relieves the suppression. In most contexts BSAP is a transcriptional activator, but in the 3′αE it inhibits the enhancer activity in at least two different ways. First, it prevents the binding of the transcriptional activator NF-αP to the αP site; second, it causes the octamer, G-rich, and κB-like motifs to exert an active repressive influence on transcription (654,656). Optimal activity of the κB site in the 3′αE may require interactions between NF-κB and proteins binding to an adjacent motif designated NFE (657).

Apart from the motifs mediating upregulation of the 3′αE during maturation to plasma cells, a response element in the enhancer for activation induced by BCR cross-linking has been traced to partially overlapping sites for the ETS family member Elf-1 and for members of the AP-1 transcription factor family (658). The same DNA sites represent a response element for CD40, though perhaps mediated by slightly different members of AP-1/ETS families (659). Two other motifs in the enhancer have been proposed to contribute to its regulation, but are less well documented: the μE1 and the μB motifs, which were first noted in the rat 3′αE and which are partially conserved in mice. The HS3 and HS4 enhancer regions of mice have been studied in less detail, but the HS4 enhancer apparently contains functional Oct-1 and BSAP binding sites (660,661).

A role for the 3′αE in isotype switching was suggested by experiments in which this region was replaced by a neomycin resistance gene (neo^r) through homologous recombination in ES cells that were then used to reconstitute the B-cell population in RAG-2 knockout mice. The resulting B cells showed normal V(D)J recombination but marked deficiencies in switching to IgG2a, IgG2b, IgG3, and IgE in vitro, whereas expression of IgM and IgG1 was normal (96). This observation suggests that the enhancer exerts isotype-specific effects on switch recombination, possibly by affecting the extent of germline transcription of the different isotypes before switch recombination. One caveat follows from the observation that, in neo^r replacement experiments to test the role of the κ enhancers, the neomycin resistance gene apparently affected κ expression beyond simple deletion of the enhancer (277,278); a similar effect of the neo^r gene may contribute to the phenotype observed in the 3′αE replacement mice.

Roles of the Two IgH Enhancers in B-Cell Development

From the experiments analyzing the Eμ and 3′αE, it would appear that Eμ functions primarily in early in B-cell development, with the 3′αE functioning later. Thus, B cells with targeted neo^r replacement of the Eμ enhancer (in chimeric RAG complementation mice) showed cis inhibition of germline transcription and VDJ recombination (662,663), whereas the similar 3′αE replacement just discussed showed normal VDJ recombination. The latter construct affected isotype switching, characteristic of late B-cell maturation. Moreover, as discussed above, spontaneous deletions of Eμ in plasmacytomas did not significantly affect Ig secretion, whereas spontaneous (646) or targeted (664) deletions removing the 3′αE depressed Ig secretion (in the latter study Eμ was also missing). Transfected and transgenic constructs driven by the 3′αE indicate that activity of this enhancer is specific to late, activated B cells (665–667), perhaps due in part to suppression by BSAP and Id3 early in the B-lineage (as discussed above) as well as to stimulation through motifs in the enhancer that are specifically responsive to antigen binding, T-cell stimuli, or mitogens such as LPS.

The κ Intron Enhancer and NF-κB

By transfecting deleted κ genes and constructs containing segments from the murine Jκ-Cκ intron linked to reporter genes, several groups demonstrated an enhancer lying about 0.7 kb 5′ from the Cκ-region gene (668,669). This location corresponds to a B-cell specific DNase I hypersensitivity site (492,493,670) and also to a segment of the intron noted to have a remarkably high degree of sequence conservation between mice, humans, and rabbits (671). The intronic κ enhancer, sometimes designated iEκ, has been dissected by fine deletions, mutations, and protein binding studies (Fig. 17C). As mentioned above, three E boxes were recognized in this enhancer, and they seem to bind to the same proteins targeted to the corresponding Eμ motifs: κE1 binds to YY1, κE2 binds to the E2A proteins, and κE3 binds to TFE3 and related proteins. E-box mutations that reduce enhancer function abolish protein binding at this motif (533).

NF-κB

An additional motif of major importance in iEκ is the κB motif GGGACTTTCC. This motif was originally discovered as a binding

site for a nuclear protein, detectable in EMSA experiments, in extracts from B cells—hence its designation NF-κB (nuclear factor-κB) (672). The presence of this protein in the same cells capable of supporting κ enhancer function provided a clue that NF-κB might be important for mediating enhancer activity. A further correlation was provided by the pre-B line 70Z/3, which has a rearranged but functionally silent κ gene. Treatment of 70Z/3 cells with LPS causes an activation of κ transcription associated with the appearance of NF-κB activity in 70Z/3 nuclear extracts (673). Mutations in the κB motif strongly reduce κ enhancer activity (533), suggesting a critical role in enhancer function. Indeed, this motif in isolation from the E-box motifs has been shown to possess enhancer activity, especially in constructs containing tandem copies of the motif (674). The enhancer activity of these constructs was much greater when transfected into a B-lymphocyte than into a fibroblast, consistent with the importance of this motif for the B-cell specificity of κ gene expression. An important physiologic role for NF-κB in mediating B-cell activation triggered by antigen recognition is suggested by the ability of surface IgM cross-linking to upregulate NF-κB activity (675). Since the discovery of the κB motif in iEκ, similar motifs have been recognized as critical functional elements regulating numerous genes outside the Ig loci. These include the genes encoding MHC class I and class II proteins and β2-microglobulin, urokinase, IL-2 and IL-2 receptor α chain, IL-6, granulocyte-macrophage colony-stimulating factor, β-interferon, inducible nitric oxide synthetase, and TNF-α and -β; the motif is also found (in tandem repeated copies) in the long terminal repeat of HIV. Many of these genes are expressed outside the B-lymphoid lineage; and indeed NF-κB was found to be inducible in T cells and HeLa by phorbol esters (673) and in other cell types by a variety of agents, including LPS, phorbol esters, TNF-α, and IL-1. Thus, rather than being B cell specific, NF-κB proteins act in many cell types, often regulated by agents associated with inflammation and often regulating inflammation-related responses. In B cells the protein regulates not only iEκ, but also the 3′αE and sites in the promoters of I regions of several H-chain isotypes, as discussed above. The complexity of NF-κB and its clinical relevance to various immune processes have inspired considerable investigation, and a large body of literature has resulted on the molecular basis of NF-κB action.

In general, the induction of NF-κB activity is not blocked by protein synthesis inhibitors, suggesting that activation must be due to modification of a preexisting protein molecule. Indeed, most cells that lack NF-κB activity in their nuclei have an inactive cytoplasmic form that is unable to bind to κB sites but can be activated when cytoplasmic extracts are treated in vitro with sodium deoxycholate (DOC) (676). DOC was found to work by abolishing the binding of an inhibitory subunit designated IκB to NF-κB (677). The active form of NF-κB initially studied is a complex of two distinct subunits of molecular weights of 50 and 65 kDa, known as p50 and p65, respectively. Thus, the initial experiments suggested a model of the inactive cytoplasmic NF-κB as a complex of p50, p65, and IκB; physiologic activators of NF-κB would abolish IκB activity, releasing the p50-p65 heterodimer to move to the nucleus, bind to κB sites, and activate transcription.

The cloning of genes encoding proteins of the NF-κB has provided a more complex and interesting picture (678). The p50 subunit and a closely related p52 protein are encoded as precursor proteins known as NF-κB1 (p105/p50) and NF-κB2 (p100/p52). The N-terminal half of both proteins share a domain of about 300 amino acids that is responsible for DNA binding. This domain is shared by an ancient family of evolutionarily conserved proteins with functionally homologous members in drosophila. The family and the domain are designated Rel, after the v-*rel* oncogene from the avian reticuloendotheliosis virus, an early recognized family member (679). The C-terminal halves of both NF-κB1 and NF-κB2 contain seven repeats of a 33-residue sequence known as the ankyrin repeat; these repeats are found in diverse proteins from bacteria to mammals and often mediate protein–protein interactions. The precursor forms of NF-κB1 and NF-κB2 are inactive and restricted to the cytoplasm because a nuclear localization sequence (NLS) is occluded by the C-terminal domain; the protein is activated by proteolytic cleavage of this domain, which uncovers the NLS of p50 (or p52). The other members of the Rel family currently known contain the Rel homology domain without an inhibitory domain; they instead contain transcriptional activation domains. These proteins include RelA (which encodes p65), c-rel (the cellular homolog of the v-*rel* oncogene), RelB, and the two drosophila proteins Dorsal and Dif. Interacting through their Rel domains, various members of the Rel family can form heterodimers or homodimers, some showing slight differences in preferred DNA binding motif. Different members function predominantly at different stages of B-cell development; in pre-B cells (and in non-B cells) p50 and p65 are seen, with p50 and c-Rel in mature B cells and LPS-treated pre-B cells, and p52 and RelB in plasmacytoma cells (657,680–682). Despite the apparent importance of the p50 subunit in B-cell development, p50 knockout mice have grossly normal κ Ig expression with a normal ratio of κ to λ L chains, possibly because of compensation by other Rel family members. However, these mice show impairments in activated Ig production (586) and in expression of certain switched isotypes, including IgG1, IgA, and IgE (114), as mentioned earlier.

Cloning of IκB genes (683–685) showed another family of proteins, all of which share the ankyrin repeats characteristic of the C-terminal half of NF-κB1 and NF-κB2 precursor proteins. This family includes IκBα, IκBβ, bcl3, and the drosophila protein cactus. Bcl3 is unusual in that when it binds to p50 or p52 homodimers it promotes nuclear localization (686) and stimulates transcription through an activation domain (687,688); this protein is not directly involved in NF-κB function in the Igκ enhancer. IκBα or IκBβ can bind to Rel family dimers, retaining them in inactive form in the cytoplasm. Many of the same stimuli that cause proteolytic removal of the ankyrin repeat-containing C-terminal half of the Rel precursor proteins also activate Rel dimer–IκB complexes through a similar mechanism. A critical step is phosphorylation of the IκB (689), apparently by an IκB-specific kinase (690,691, 691a). The phosphorylated protein is a target for addition of the small protein ubiquitin, a modification that flags proteins for destruction by proteasomes (692). IκBα and IκBβ show somewhat different preferences for specific Rel dimers, different tissue distribution, and an interesting difference in regulation. The IκBα promoter contains multiple κB sites, which are activated when IκB destruction releases NF-κB, allowing this protein to migrate to the nucleus and stimulate IκBα resynthesis, which causes inhibition of NF-κB. Thus, in cells where IκBα predominates, NF-κB activation is short lived (693). In contrast, IκBβ is not regulated in this way, so in cells where it predominates, NF-κB activation may be more prolonged (685). The importance of NF-κB for κ expression was supported by transfection of AMuLV pre-B cell lines with an engineered form of IκB capable of suppressing both RelA and c-Rel; this dual block markedly inhibited germline κ transcription and rearrangement (694).

A MAR and Silencer Elements Upstream from iEκ

Upstream from iEκ lies an A/T-rich region that has been identified as a MAR (695). Kappa gene constructs inserted into the mouse genome as transgenes or stably integrated into B cells demonstrated that this MAR contributes to transcription of the associated gene (696,697) and to associated demethylation (698). [The demethylation also appears to be regulated by NF-κB (699).] Within the MAR an AP1 binding site has been reported (700). This site appears to be required for optimal enhancer activity when transfected into LPS-treated pre-B cells and mature B cells and shows LPS-inducible binding in EMSA experiments. In transfections into HeLa and T cells, a 232-bp fragment 5′ from the κB sequence inhibited expression of a linked gene, whereas this fragment did not affect expression in B cells (701). Thus, this fragment may contain a gene silencer that is active in non-B cells and contributes to the B-cell specificity of the enhancer. A shorter element designated κNE (negative element) lying just upstream from κB, and conserved at this position in humans and rabbits, has been reported to inhibit enhancer activity (702); this inhibition was reversed in a B cell–specific manner by another element a few base pairs upstream.

The 3′ κ Enhancer

Components of the Enhancer

As in the H-chain system, the search for a second enhancer downstream from the Cκ gene was inspired by a cell line whose expression of its endogenous κ gene seemed difficult to explain in terms of the known intronic enhancer. Thus, the myeloma S107 was found unable to support transcription of transfected constructs driven by the κ intronic enhancer because this line lacks NF-κB activity, yet it is able to transcribe its endogenous κ genes (703). A search 3′ from the Cκ gene showed a second enhancer about 9 kb downstream from Cκ (704). This enhancer is about sevenfold stronger than the κ intronic enhancer and in transfection experiments is B cell specific. Inclusion of the 3′ enhancer in transgenic constructs leads to more than 20-fold higher transgene expression than observed with constructs lacking this sequence (705). Like the iEκ, the 3′Eκ can be activated in pre-B cells by LPS. The functional elements of the murine enhancer have been dissected, demonstrating a complex set of motifs mostly clustered in a 132-bp core enhancer (Fig. 19D). One important motif identified by deletion and multimerization constructs contains the sequence CATCTGTT, which conforms to the CANNTG consensus for HLH binding motifs; indeed, this motif appears to bind to such a protein because the activity of multimerized versions of this motif, as well as the activity of the entire enhancer, can be inhibited by the HLH protein Id described above (706). The second principal motif binds PU.1, a member of the ETS domain family of transcription factors described earlier in this section. The binding of PU.1 to the enhancer recruits a second B cell–specific protein—designated PU.1 interaction protein or Pip (formerly NF-EM5)—which binds to an adjacent DNA segment whose integrity is necessary for full enhancer function (707,708). Pip is homologous to members of the interferon regulatory factor (IRF) family of transcription factors, and its binding to PU.1 is also important in the murine λ enhancers described below. The Pip-PU.1 interaction requires phosphorylation on a particular serine residue of PU.1, suggesting the possibility that the degree of phosphorylation might contribute to physiologic regulation of enhancer activity.

Upstream from these two motifs lies a sequence that was found by mutation analysis to be necessary for maximal enhancer activity and which was found, by λgt11 library screening, to bind to the transcription factors ATF-1 (activating transcription factor) and CREM (cyclic AMP response element modulator) (709). Both of these proteins can bind to PU.1 in vitro. The ability of CREM to function in the enhancer is supported by the observation that dibutyryl cAMP can increase the 3′Eκ enhancer activity. The DNA binding motif for these factors also corresponds to an AP-1 site, and the components of AP-1, c-Fos, and c-Jun were also found to activate the enhancer through this motif (710). Indeed these two AP-1 subunits can participate in a higher order complex with PU.1 and Pip that is detectable biochemically by EMSA and functionally by synergistic activation of the enhancer by these proteins when expressed in fibroblasts, in which the enhancer is normally silent. A site detected by in vivo footprinting that is occupied in pre-B and B cells, but not in plasma cells, was shown in EMSA experiments to bind a protein with characteristics of BSAP (711). At about 90 bp upstream of the core enhancer is an additional site that binds the transcription factor Sp1 and that is required for maximal enhancer activity in some constructs (712).

Just downstream from the murine 3′ κ enhancer is a negative regulatory region that seems to suppress the activity of this enhancer in pre-B cells (626). One component of this region appears to be a binding site for the zinc-finger protein YY1 or NF-E1 discussed earlier in the context of this protein's function in Eμ. An additional component may be a binding site for lymphocyte enhancer factor-1 (LEF-1), an HMG-related protein that binds in the minor groove of DNA and causes DNA bending; this protein is also a known component of TCR-α enhancer regulation. Binding activity at the presumptive LEF-1 site was depressed by treatment with LPS, suggesting the possibility that this effect may explain how the 3′Eκ enhancer activity might be upregulated by LPS in the absence of a site for NF-κB (713).

Roles of the Two κ Enhancers in B-Cell Development

The intronic κ enhancer seems critical for supporting Vκ-Jκ recombination because targeted disruption of the enhancer by neo^r replacement severely impairs or abolishes such recombination (277,714). Based on transfection experiments suggest that this enhancer is moderately active in pre-B cells but can be upregulated with activating agents such as LPS and phorbol esters, which exert their effects through NF-κB activation. Although the enhancer is active in plasmacytoma cells, its integrity does not seem critical for κ gene expression at this stage, when the 3′Eκ is more active; however, some transfection experiments suggest that the two enhancers may function synergistically in mature B and plasma cell stages (715,716). Even at the pre-B stage the 3′Eκ is active and can support lineage-specific expression of a transgenic κ gene lacking iEκ (717). Both enhancers seemed necessary to support somatic mutation in κ transgenes (432), although some of this effect may be mediated by effects on transcription. In mice with targeted disruption of 3′Eκ, decreased numbers of κ-expressing B cells were observed, as though this enhancer contributes to Vκ-Jκ recombination (278), although in transgenic animals harboring constructs capable of Vκ-Jκ joining, the 3′Eκ seemed primarily to suppress recombination. In the absence of this enhancer, recombination occurred in T cells or prematurely in pro-B cells (264). In vivo footprinting studies have shown that changes in activity of the two enhancers during B-lineage develop-

ment are accompanied by changes in occupancy of specific motifs by nuclear binding proteins (711,718).

λ Enhancers

For many years the λ locus frustrated investigators searching for an enhancer in the Jλ-Cλ intron, but with the recognition of enhancers downstream from C genes, attention turned to these regions and λ enhancers were identified (719). Highly homologous B cell–specific enhancers are located 15.5 kb downstream from murine Cλ4 (the $E_{\lambda 2\text{-}4}$ enhancer) and 35 kb downstream from Cλ1 (the $E_{\lambda 3\text{-}1}$ enhancer). Four functional motifs were identified in each enhancer (720). Two, λA and λB, are critical for enhancer function in that mutations in either abolish activity, although neither of them are active when present in multimerized form in constructs. The λB motif binds to PU.1 and Pip (721,722), which also bind together in the 3′Eκ, as discussed above. The λA and λB elements are flanked by E box–like motifs, which may bind to HLH proteins active in other Ig enhancers.

Human Ig Enhancers

The above account of regulatory regions in the three Ig loci has focused on the mouse genes because these regions were discovered before homologous regions of other species, and they have been most thoroughly studied. Sequence analysis of homologous human regions have shown a high degree of conservation, especially in the core enhancers containing functional transcription factor binding motifs; and other enhancer properties—including DNase hypersensitivity, in vitro protein binding, and functional enhancer activity—also have been documented for human Ig enhancers. Some differences in enhancer number have resulted from gene duplications specific to mice or humans. Thus, in contrast to the two murine λ enhancers, the human λ locus includes a single enhancer downstream from Cλ7 (723–725). Conversely, the duplication of the two γ-γ-ε-α segments of the human IgH locus has led to duplicated enhancer complexes downstream from the human Cα genes (491,726). Analyses also have been conducted on the human iEκ (493,671,727–729), the human 3′Eκ (140,730), and the human Eμ (486,731–734).

Generalizations Concerning Ig Transcriptional Regulation

Each gene segment within the three Ig loci is regulated by nearby DNA regions outside the coding exons. The regulatory regions are composed of several motifs that control transcription by binding to specific nuclear factors that can stimulate or inhibit transcriptional initiation. Some of the motifs are shared between different enhancers or promoters and some are unique to one region. The presence (or activity) of the nuclear factors in different cell types often correlates with the expression of the associated gene segment. The research to date seems to have identified many components of a complex regulatory machinery, but there are many gaps in our understanding. We know little about what regulates the nuclear factors—how the known components change in response to cell maturation and to external signals such as antigens, cytokines and T cells—nor do we understand how the actions of these nuclear factors are integrated with other chromosomal changes such as histone acetylation, DNA methylation, matrix attachment, and nucleosome repositioning.

APPLIED SCIENCE OF IMMUNOGLOBULIN GENES

Up to this point, this chapter has considered how our current knowledge of Ig genes can explain the antibody response. In the present section we briefly address several examples of other areas to which this knowledge has been applied with interesting results.

Ig Genes in Lymphoid Malignancies

Many malignant tumors have been shown to derive from single transformed cells that have undergone clonal expansion with failure of normal cellular controls. Lymphoid malignancies of the B-lineage provide a classic demonstration of clonality because they derive from cells with unique genetic material (rearranged Ig genes) distinct from the bulk DNA of the same organism. Similarly, analyses of TCR genes have been valuable in establishing clonality of T-cell malignancies. Examination of both gene systems is useful in establishing the lineage of neoplasms that lack characteristic phenotypic markers and in detecting clonal rearrangements as a marker for malignancy; a huge clinical literature has accumulated (735–738).

The first general strategy to analyze clonality has been to isolate genomic DNA from neoplastic and normal tissue from the same patient and to examine it for Ig or TCR gene rearrangements using Southern blotting. A nonclonal population of B-lymphocytes contains a mixture of rearrangements so numerous that, after restriction enzyme digestion and Southern analysis, the rearranged fragments bearing the JH sequence are spread thinly throughout the length of the gel, so no specific rearranged band is detectable. In contrast, clonal expansions of a specific rearrangment as in a lymphoid malignancy will produce a distinct rearranged band, detectable even when the malignant cells are present as a minority in the cell population. Admixture experiments have demonstrated that rearranged bands can be detected when malignant clonal cells represent as little as 1% of the population, although 5% is a more typical detection limit. The Southern blotting technique can be used on peripheral lymphocytes as well as on biopsy specimens of solid tumors. More recently the power of the PCR to amplify miniscule amounts of DNA has made it possible to assess clonal rearrangements from samples as small as histologic tissue sections or to detect rare leukemic cells in a 10^5 excess of normal cells; such assessments are critical for clinically important judgments (739). Ig gene rearrangements have been demonstrated in acute lymphoblastic leukemia (ALL), chronic lymphocytic leukemia (CLL), multiple myeloma, B-cell follicular and diffuse lymphoma, hairy cell leukemia, B-cell prolymphocytic leukemia, Hodgkin's Disease, Burkitt's lymphoma, and in the blast crisis of chronic myelogenous leukemia (738).

DNA analysis has been particularly revealing in the case of ALL cells, which before the advent of DNA analysis were generally (80%) of uncertain lineage, lacking both B-and T-phenotypic markers. The majority of these null ALL samples analyzed contain rearrangements of H-chain Ig genes, and about 40% also have rearranged L-chain genes, although no surface Ig is detectable. The cells thus typically resemble the pre-B stage of lymphoid development. During the course of the disease some ALL cells show clonal evolution of additional Ig gene rearrangements, thus further matching the pre-B phenotype. When DNA from serial peripheral blood samples of pre-B ALL patients are examined by Southern blotting, the clonal rearranged band can be used as a marker of leukemic remission and relapse (740,741). Ig gene rearrangements also have

been sought in T-cell ALL; rearrangement of the Ig H-chain gene occurs rarely in these malignancies, and L-chain rearrangement is not observed. Conversely, TCR gene rearrangement occurs in roughly half of pre-B cell ALL, especially at the TCR-γ locus (740). These examples of lineage infidelity may reflect a developmental stage before definitive commitment to B- or T-differentiation when both gene systems may be susceptible to the recombinase machinery.

Chromosomal Translocations Involving Ig Gene Loci

In contrast to the rearrangements in tumors discussed so far, which represent physiologic recombination events that occurred in premalignant progenitor cells, an entirely different kind of Ig gene rearrangement has been found in several lymphoid neoplasms: rearrangements that appear to have played a role in the malignant transformation itself. All except the two most recently discovered cases involve genes normally expressed in the B-cell lineage.

c-myc *Translocations in Burkitt's Lymphoma*

The first example to be elucidated—and one that still represents a prototype—was in Burkitt's lymphoma, where a consistent pattern of chromosomal translocation has been observed involving a reciprocal exchange between chromosome 8 and either chromosome 14, chromosome 2, or chromosome 22. The latter three chromosomes contain the three human Ig gene loci (H chain, κ, and λ) (47,742), and the Ig genes were mapped by in situ hybridization to the same bands involved in the chromosome 8 translocations. The translocation breakpoints have been cloned and sequenced, providing a detailed picture of these nonphysiologic recombination products (Fig. 20).

The sequence consistently donated from chromosome 8 has been found to be the c-*myc* oncogene, the mammalian cellular homolog of the oncogene first identified in avian leukosis virus (743). Translocation of c-*myc* into the IgH locus is also observed in the 12;15 translocations commonly seen in murine plasmacytomas. The CH regions most frequently involved are α (in the mouse) and μ (in human cells), generally the alleles on the nonexpressed chromosome. These translocations leave the IgH and c-*myc* genes joined head to head (in opposite orientations). As shown in Fig. 20, the first exon of the c-*myc* gene is commonly absent from the IgH-associated translocation product, but because this is a noncoding exon, such genes can still encode a functional protein. The site of the translocation can vary over a wide distance for both the c-*myc* and IgH genes, which may be separated by more than 100 kb in some 14q+ chromosomes. Generally the translocations in Burkitt's lines fall into two categories roughly paralleling the two clinical forms of the disease: the endemic African type and the sporadic type (744).The endemic Burkitt lines seem to represent an early B lymphoid stage as they make primarily membrane Ig; these lines demonstrate Ig locus breakpoints near V and J regions and c-myc breakpoints far 5' of exon 1. However, recent reports that endemic lines show evidence of somatic mutation (744a,744b), combined with the realization that VDJ recombination may be reactivated in the GC (as discussed earlier), have suggested that endemic lines may represent later stages in B-cell development than investigators initially thought. The less frequent translocations involving κ and λ bring these genes downstream from c-*myc* and oriented in the same 5′-3′ direction, as shown in Fig. 20. Analysis of these translocations has incidentally provided an assignment of the 5′-3′ orientation of the normal Ig gene loci with respect to the centromeres on their respective chromosomes.

Several observations suggest that the translocation event bringing the c-*myc* gene adjacent to the Ig gene locus participates in the malignant transformation of the progenitor lymphocyte. In normal cells the Myc protein plays a complex role in regulating cell cycle progression, probably through transcriptional activation of genes associated with cell division. The Myc protein has a structure typical for the bHLH-zip family of HLH transcription factors discussed earlier in this chapter. In association with a heterodimer partner named Max, Myc binds in vitro to a typical E-box motif CACGTG (745) and regulates several genes that might be relevant to its role in regulating proliferation (746,747). In Burkitt's lines transcription is generally maintained at relatively high steady-state levels, which could contribute to the malignant proliferation of these cells. Among somatic cell hybrids constructed between a Burkitt's lymphoma and mouse myeloma, human c-*myc* transcripts were found in hybrid cells containing the 14q+ chromosome but not in those with the normal chromosome 8; this finding suggests that translocation of the c-*myc* gene into the Ig locus was responsible for activating its transcription in cis (748). This activation could be mediated by proximity to the intronic μ enhancer in some Burkitt's lymphomas. In others, in which this intronic enhancer is absent from the translocated c-*myc* locus, c-*myc* activation might be mediated by the 3′αE and associated regulatory regions, which have been shown to be competent to stimulate transcription from the c-*myc* promoter (490). The potential long-distance regulation implied by the LCR properties of these regions may explain the dysregulation of c-*myc* expression in some Burkitt's lines even where the c-*myc* gene is relatively distant from the IgH locus. The observation of T-cell leukemias with c-*myc* translocations into the TCR-α chain locus (749) support the generality of deregulated c-*myc* expression in oncogenesis.

Other Oncogenic Translocations Involving the Immunoglobulin Loci

In addition to the translocations of the c-*myc* gene on chromosome 8, other translocations involving the IgH locus have been reported in lymphoid malignancies, and translocation breakpoints have been cloned in the hope of identifying new protooncogenes that—by analogy with c-*myc*–might be activated by the translocation. An 11;14 translocation seen in some CLLs and centrocytic B-cell lymphomas was found to join the nonexpressed IgH locus to a region of chromosome 11 that has been termed *bcl*-1 (B-cell leukemia/lymphoma-1) (750). Although attempts to detect a deregulated transcription unit in this region were initially unsuccessful, an oncogene candidate has emerged that was first discovered as a partner in a different chromosomal rearrangement, one involving the parathyroid gene (751). This oncogene, known as PRAD-1 (parathyroid adenomatosis-1) encodes cyclin D1—a regulator of cell division—and maps to the same band (11q13) involved in the translocations with the IgH locus. Cyclin D1/PRAD-1/Bcl-1 transcripts are elevated in several CLL lines with *bcl*-1 translocations (in contrast to other CLLs lacking this translocation) and in approximately 90% of mantle-cell lymphomas (752).

Another translocation involving chromosome 14 occurs in the majority of cases of follicular lymphoma and involves chromo-

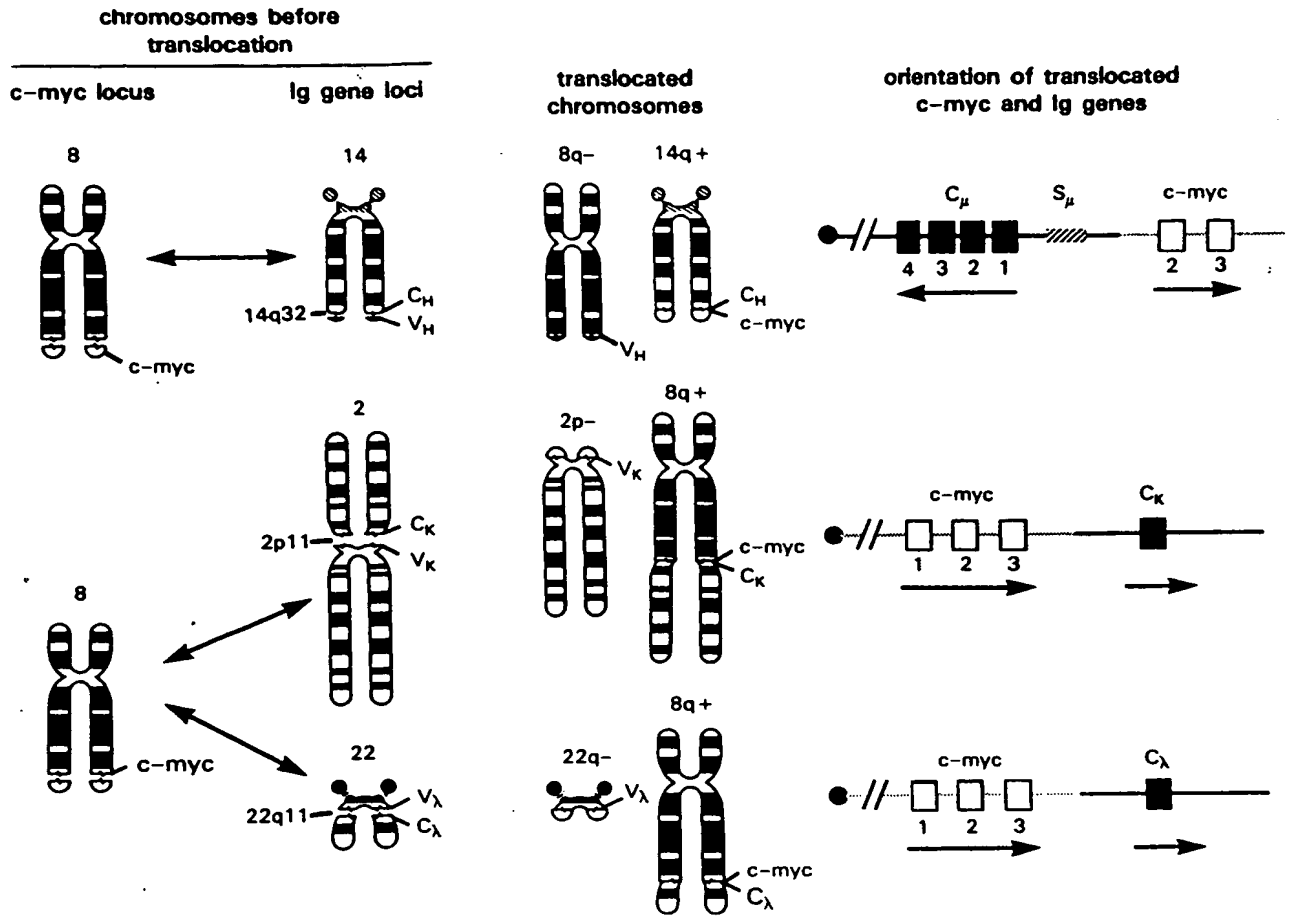


FIG. 20. Translocations between Ig genes and the c-*myc* locus observed in lymphoid malignancies. The c-*myc* gene on chromosome 8 (*left*) is associated with translocations resulting from chromosomal breaks either above (*top*) or below (*bottom*) the gene itself. In the former case, translocation to the IgH locus at 14q32 leads to a 14q+ chromosome bearing c-*myc* sequence and IgH sequence in opposite orientations (*upper right*). In the latter case translocations with the Igκ locus at 2p11 or the Igl locus at 22q11 lead to junctions in which the *myc* and Ig genes are in the same orientation.

some 18 band q21. Analysis of cloned fragments containing the translocation breakpoint led to the identification of a new oncogene designated *bcl*-2 (753). The 14;18 translocation creates a *bcl*-2–Ig fusion gene whose mRNA transcripts are elevated as a result of both transcriptional deregulation and altered RNA processing. The *bcl*-2 gene is unusual among protooncogenes in that its normal role appears not to be the promotion of cell proliferation but rather the inhibition of programmed cell death or apoptosis (754). It was the first identified member of a family of proteins that form heterodimers that stimulate or inhibit apoptosis (755). The Bcl-2 protein is expressed in tissues where some cell populations undergo apoptosis but selected subsets are spared (756). In particular, the protein can be detected in the apical light zone of GCs, where it is believed that somatic mutation of Ig genes is ongoing and only the subset of B cells displaying surface Ig with high binding affinity for antigen can survive. Bcl-2 protein is also expressed in surviving T cells in the thymic medulla. Transgenic mice with a *bcl*-2–Ig minigene developed a lymphoproliferative syndrome due to the extended life span of their lymphoid lineage; eventually most mice progressed to a lymphoma (757), supporting an important role of the *bcl*-2–Ig translocation in malignant transformation.

A translocation involving the IgH locus and chromosome 19 band q13.1 is a recurring but uncommon abnormality in CLL (758,759). The cloned translocation breakpoint showed a gene, designated *bcl*-3 (760), that encodes a protein containing seven of the ankyrin repeats characteristic of IκB-like proteins, and which was discussed earlier in this chapter. The recombination is often head to head near Sα in the IgH locus and leads to a marked elevation of intact Bcl-3 transcripts in CLL lines with the 14;19 translocation, as compared with lines without this abnormality. The identity of the genes regulated by Bcl-3 is not known, although presumably these genes are normally regulated by Rel family dimers.

Bcl-6 was first identified by cloning the breakpoint at the most common translocation observed in B-cell non-Hodgkin's lymphoma, that involving the IgH locus and 3q27 (761). Intact bcl-6 transcripts are increased by the translocation. The 95-kDa DNA-

binding Bcl-6 protein has six zinc finger domains and is able to mediate strong transcriptional repression, largely as a result of its N-terminal POZ domain (762). Bcl-6 transcription is downregulated on B-cell activation, but the protein is detected at relatively high levels in GCs (763). Bcl-6 knockout mice fail to form GCs and do not show affinity maturation, but have diffuse inflammation with an increase in IgE-bearing lymphocytes (764,765).

The bcl-7A gene at 12q24.1 (766) and *bcl*-8 gene at 15q11-13 (767) were cloned from breakpoints in IgH translocations in lymphomas. The Bcl-7A sequence appears homologous to the actin-binding protein caldesmon, and the *bcl*-8 gene is expressed in testis and prostate; but little more is known about the function of these genes at present.

Many of the translocations described above can be detected by PCR as clonally unique amplification products, which can be used as markers for minimal residual disease (768).

Hybrid Recombinations

Another class of aberrant chromosomal rearrangements involving the IgH locus includes the chromosome 14 inversions observed in some T-cell lymphomas (769,770). These remarkable rearrangements occur between the IgH locus at 14q23 and the TCR-α locus at 14q11 and clearly appear to have been mediated by the V assembly recombinase. In one well-studied example, an Ig VH segment is joined to a TCR Jα segment on the telomeric end of the chromosome, whereas in the centromeric region a signal joint is found; because this joint is not reciprocal to the VH-Jα coding joint on the same chromosome, at least two recombination steps must have occurred. In these chromosomal inversions no oncogene sequence seems to be involved, so their relationship to the malignancy is uncertain. By PCR, similar hybrid antigen receptor recombinations can be detected at low levels in normal individuals; these recombinations are present in higher than normal frequencies in patients with ataxia telangectasia and in agricultural workers exposed to chemicals, a population with increased risk for lymphoid malignancy (771).

Abnormal Ig Gene Loci in Disease

Several immune deficiency diseases are associated with selective or global decrease in serum Ig levels. Although it might have been expected that elucidation of the Ig gene loci would clarify the molecular basis for these diseases, very few examples of genetic defects in Ig genes have been reported. Indeed, examination of restriction fragment length polymorphisms (RFLPs) detected with probes in the IgH locus have indicated that at least two genetic defects associated with Ig H chains (familial selective IgA deficiency and the hyper-IgM syndrome) are not linked to the IgH genes. Many Ig deficiencies are undoubtedly caused by cellular abnormalities in the complex mechanisms of B-cell development, T-cell interaction, lymphokine response, antigen triggering, and so forth.

In one of the rare examples of an Ig deficiency due to Ig gene mutation, defects in the Cκ genes have been described in a patient with selective deficiency of κ synthesis (772). Different point mutations in each of the Cκ alleles of a patient were observed, leading to amino acid replacements that could have disturbed the intradomain disulfide bonds critical to Ig structure.

In the H-chain locus six large deletions have been described in humans (Fig. 7), the largest involving the loss of the γ1, ψε1, α1, γ2, and γ4 genes (55); despite the complete absence of the corresponding H chains in the serum, individuals with homozygous deletions generally show no clinical evidence of immunodeficiency. However, homozygous defects in the μ gene can result in agammaglobulinemia (772a).

The contribution that specific polymorphic V genes might make to autoimmune disease has been explored by several investigators. Although human V gene sequences are remarkably well conserved between individuals, it is known that human V loci display large insertion/deletion polymorphisms in the population (381,382, 773–777). Extrapolating from this fact, it might be supposed that the presence of specific unusual germline V genes could increase the risk for certain autoimmune responses, much as certain MHC haplotypes are associated with increased risk for such diseases. However, the genes expressed in autoimmune antibodies have not proved to be rare in the population; and although some disease associations with V haplotypes have been reported (778,779), genetic variation in V genes does not appear to be a major risk factor. Specific immune defects due to the absence of specific V regions are also possible. In Navajos a defective copy of the VκA2 gene—which encodes the predominant Vκ chain in antibodies against *Haemophilus influenzae*—has been suggested as possibly contributing to the high susceptibility in this population to infections with this bacterium (780).

Genetic Engineering of Ig Genes

Using the considerable knowledge of Ig genes that has been gained from modern molecular cloning techniques, a number of investigators have been exploiting these genes as bioengineering tools for various basic research and applied science goals. Although a detailed treatment of these studies is beyond the scope of this chapter, we will briefly consider a few of the more interesting ideas.

One basic research goal is the exploration of structure–function relationships of Ig molecules by engineered modifications of Ig structure. The IgM molecule has been studied in this way initially by exploiting natural mutant hybridoma lines making abnormal antibodies. The abnormal μ genes were cloned and sequenced, and observed mutations that were candidates for causing the phenotypic abnormality were either reverted or reintroduced into normal genes by site-directed mutagenesis to verify their effects. Using this approach, a 39-bp deletion near the C-terminal end of the molecule was found to prevent pentamer formation (781) and replacement at codon 436 in the CH3 domain was found to depress complement-activated cytolysis (782). Sequences responsible for differential complement activation by human γ isotypes were identified by exchanging various residues from one isotype to another and following the resulting effects on complement activation (783). In an analysis of the mouse γ2b H chain, systematic alteration of amino acids on the surface of the CH2 domain by in vitro mutagenesis led to the identification of three residues critical for the binding of the complement factor Clq (784). A structure–function analysis of the ε H chain was undertaken by testing ε chain fragments—generated by bacterial expression of engineered segments of the ε gene sequence—for their biologic activity. A 76–amino acid fragment spanning the CH2-CH3 boundary was found to bind to mast cells in vitro and to inhibit the action of IgE *in vivo* (Prausnitz-Kustner reaction) (785). For V-region structure–function analysis, site-directed mutagenesis of V regions has been used to study the determinants of antibody affinity and specificity (786). These studies suggest that biotechnology can provide powerful methods for analyzing important features of Ig protein structure.

When structure–function relationships are sufficiently understood, the next logical engineering challenge is to improve on nature, designing Ig molecules with specific desired properties by modifying appropriate segments of Ig genes. One goal has been to combine the advantages of human and murine monoclonal antibodies to make medically useful products. Murine hybridomas grow quickly, produce large amounts of antibody and are quite stable relative to human hybridomas, which are generally poor in all three respects. Yet for many applications—like the use of antitumor antibodies in human patients—the more easily generated mouse monoclonals would be unsatisfactory because of their immunogenicity and their relative inefficiency in generating C region–dependent effector functions (such as complement fixation and antibody-dependent cellular cytotoxicity). A solution that has been tested by several laboratories is to construct chimeric genes linking a human C-region gene to a murine V region cloned from a mouse hybridoma generated against the antigen of interest; these constructs are then transfected into nonsecreting variants of mouse hybridomas yielding transfectomas that secrete humanized antibodies with murine V regions and human C regions (787–789). To reduce immunogenicity arising from the murine V-region sequences, the CDRs from a murine antibody of desired specificity can be grafted onto human V-region framework sequences (790,791). A completely different approach to obtain human antibodies using murine hybridoma technology involves engineering mice to express human antibodies (792). This ambitious goal was achieved starting with mice whose endogenous Ig κ and IgH gene loci were disrupted by homologous recombination. ES cells from these mice were then fused with yeast spheroplasts containing YAC constructs of human DNA. The resulting "xenomice" bear 66 VH regions, about 80% of the human VH repertoire, and the complete DH, JH, μ, δ, and γ2 C-region genes, including intronic and 3′α enhancers. The transferred κ locus contains most of the proximal part of the Vκ locus (32 Vκ genes), Jκs, and Cκ as well as both the intronic and 3′ enhancers and the kde. The human genes support grossly normal development of murine B cells, most of which secrete exclusively human Ig; about 15% of B cells express human H chain with mouse λ chain. The human gene loci support antigen-specific antibody responses to immunization, demonstrating isotype switching and somatic mutation. Xenomice can be used to generate hybridomas that secrete human Ig but that provide all the advantages associated with their murine origin.

As further variations on Ig structure, bioengineers have designed antigen-binding molecules that do not require combining separate proteins containing the L-chain and H-chain V regions. One approach is to exploit camel Igs, which lack L chains but still can bind antigens efficiently (793). A more widely studied strategy involves single-chain Fv proteins (794,795); these can be obtained from gene constructs encoding hybridoma-derived Vκ and VH domains connected by a flexible synthetic linker of about 15 amino acids that allows these two domains to associate via the same protein–protein interactions that hold them together in a normal antibody. The two domains thus form a composite antigen-binding structure that often retains the specificity and affinity of its parent monoclonal antibody.

The paradigm has facilitated another engineering advance: a scheme for generating monoclonal antibodies without hybridoma fusions. Libraries of amplified Vκ and VH regions are cloned together into the same filamentous phage vector, which is designed to express both V regions as an Fv fusion protein on the outer surface of the phage. Such phage display libraries can then be selected for antigen binding by passage over an antigen-containing affinity column (796) or by successive precipitations with antigen (797). Phage clones selected for antigen binding contain the Vκ and VH genes encoding an effective antigen-binding domain. In several libraries of 10^6 clones, antigen-binding phage could be found only if the V-region sequences were derived from B cells that were obtained after immunization, but for considerably larger libraries, prior immunization is not necessary. Once an antigen-binding clone is obtained, it can be subjected to random mutagenesis and further rounds of selection to obtain higher affinity antigen binding (798). Mutagenesis can be achieved by chemical mutagens, by error prone PCR, by shuffling Vκ and VH chains between constructs, by passage through a mutator strain of bacteria or by a strategy of codon-based mutagenesis (799). A dramatically effective mutagenesis strategy with great promise for exploring the sequence space of antibody structures allows shuffling of mutations at different positions within a protein sequence to assemble various combinations of mutations before selection (800). Once a combination of high-affinity Vκ and VH have been selected, the individual Vκ and VH regions can be subcloned and inserted back into appropriate expression vectors to generate Ig molecules (801,802). It is debatable whether antibody sequences obtained by the phage display library strategy are typical of natural antibodies, but it is clear that the technology has general utility for producing high-affinity monoclonal antibodies of various specificities without hybridomas.

Bioengineering technologies have been used to alter the natural sequences of both C and V regions to obtain proteins with particular properties. Tinkering with C-region sequences can improve function of engineered antibodies. For example, in the CH3 domain of a γ1 monoclonal, replacement of a serine residue by cysteine led to dimerization and a dramatic improvement in function of an antileukemia antibody (803). One particularly interesting use of V-region engineering is the design of antibodies with catalytic activity. Enzymes are thought to catalyze reactions in part by reducing the activation energy—that is, stabilizing activated transition-state intermediates by strong binding interactions. Several groups have shown that, by a similar mechanism, antibodies directed against a molecule resembling the transition state of a chemical reaction can catalyze that reaction (804,805). Catalysis also can be achieved by Fv proteins (806). Site-directed mutations of V-region sequences can be used to analyze and enhance the catalytic activity (807). Indeed, if the antibody catalysis can be engineered to replace an in vivo loss mutation of an essential enzyme, random mutations of the antibody gene can be selected in vivo for improved activity (808).

In another avenue of Ig engineering, antibody domains have been added to unrelated peptide sequences in order to confer some desired Ig function to a different protein. Most commonly the V region is used to direct the unrelated polypeptide to a specific target. For example, in an attempt to improve the potency and specificity of tissue plasminogen activator (t-PA, an enzyme useful in dissolving clots in heart attack victims) an antifibrin antibody was linked to t-PA sequence to focus the plasminogen activation on fibrin clots (809). Other similar engineering projects have linked staphylococcal nuclease and *Escherichia coli* DNA polymerase functions to Ig molecules. A major area of applied research explores Ig-toxin hybrids, which offer the potential of delivering potent toxins to specific targets, especially cancer cells (810). A related strategy is to link a V region of one specificity to an Ig of a second specificity, creating bispecific antibodies or diabodies. Bis-

pecific antibodies have a wide range of uses, including targeting T cells (via anti-CD3) to cells bearing a particular antigen that can be recognized by an antibody rather than by a TCR (811); they also can be used in immunoassays (812). Bispecific antibodies can be generated by transfecting the two L-chain and two H-chain genes into a single producer cell, or by fusing cells producing two antibodies; in either case the desired bispecific protein must then be purified from the resultant mixture of components. Alternatively, bispecific antibodies can be genetically engineered as two Fv proteins joined by a linker chain of amino acids (813).

Although the above examples represent uses of Ig V regions, Ig C regions also have been exploited, often fused to unrelated proteins that have their own targeting properties; such fusion proteins are known as immunoadhesins. In such constructs the Ig constant domains can confer multivalency, increased stability, and effector functions (e.g., binding to Fc receptors) that can be useful for certain applications (814). For example, the receptor for most human rhinoviruses is ICAM-1, and a soluble form of this protein might act as a decoy receptor to block infection. An ICAM-1 molecule fused to Ig H-chain domains was a much more efficient inhibitor of infection than ICAM-1 alone (815).

Antibody gene constructs have been expressed in a variety of production systems, including B-lymphoid lines (transfectomas), other mammalian lines (e.g., COS cells), and bacteria. More recent experiments have involved expression in insect cells (816) and in plants (817), where they could theoretically be ingested in an unpurified state to confer passive mucosal immunotherapy. Ig gene constructs also have been designed so that the antibody is not secreted, but instead binds (as an "intrabody") to intracellular targets (818).

The strategies described above have used the coding sequences of Ig genes, but regulatory sequences also have been used in bioengineering projects, primarily to obtain B cell–specific expression of foreign gene constructs introduced nonspecifically into multiple cell types. Transgenes, for example, are present in every cell type, but transgenic constructs linking the μ enhancer to the c-*myc* gene have induced malignancies specific to the B-cell lineage, in which this enhancer is active (819,820). By similar logic, a retroviral construct containing an intracellular toxin (like diphtheria toxin) programmed for B cell–specific expression might be used to treat B-cell malignancies (821).

CONCLUSION

Recombinant DNA technology has revolutionized the study of the antibody response. Initial investigations used powerful cloning and sequencing methods to define the structure of the Ig genes as they exist in the germline and in actively secreting B-lymphocytes. More recent studies have probed the mechanisms of the processes unique to these genes, i.e., rearrangements and somatic mutation. These more difficult questions will represent a challenge for a long time to come, although the recent experiments yielding VDJ and switch recombination reactions in cell-free extracts can be expected to yield valuable clues to the mechanisms of these processes. Meanwhile the knowledge already gained about Ig genes is being applied to many clinical and scientific endeavors that hold promise for exciting advances in the near future.

REFERENCES

1. Dreyer WJ, Bennett JC. The molecular basis of antibody formation. *Proc Natl Acad Sci USA* 1965;54:864–869.
2. Brack C, Hirama M, Lenhard SR, Tonegawa S. A complete immunoglobulin gene is created by somatic recombination. *Cell* 1978;15:1–14.
3. Rabbitts TH. Evidence for splicing of interrupted immunoglobulin variable and constant region sequences in nuclear RNA. *Nature* 1978;275:291–296.
4. Seidman JG, Max EE, Leder P. A kappa-immunoglobulin gene is formed by site-specific recombination without further somatic mutation. *Nature* 1979;280: 370–375.
5. Max EE, Seidman JG, Leder P. Sequences of five potential recombination sites encoded close to an immunoglobulin kappa constant region gene. *Proc Natl Acad Sci USA* 1979;76:3450–3454.
6. Sakano H, Huppi K, Heinrich G, Tonegawa S. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature* 1979;280:288–294.
7. Tonegawa S, Maxam AM, Tizard R, Bernard O, Gilbert W. Sequence of a mouse germ-line gene for a variable region of an immunoglobulin light chain. *Proc Natl Acad Sci USA* 1978;75:1485–1489.
8. Parhami SB, Margolies MN. Contribution of heavy chain junctional amino acid diversity to antibody affinity among p-azophenylarsonate-specific antibodies. *J Immunol* 1996;157:2066–2072.
9. Fang W, Mueller DL, Pennell CA, et al. Frequent aberrant immunoglobulin gene rearrangements in pro-B cells revealed by a bcl-xL transgene. *Immunity* 1996;4: 291–299.
10. Gu H, Kitamura D, Rajewsky K. B cell development regulated by gene rearrangement: Arrest of maturation by membrane-bound D mu protein and selection of DH element reading frames. *Cell* 1991;65:47–54.
11. Corbett S, Tomlinson I, Sonnhammer E, Buck D, Winter G. Sequence of the human immunoglobulin diversity (D) segment locus: A systematic analysis provides no evidence for the use of DIR segments, inverted D segments, "minor" D segments or D-D recombination. *J Mol Biol* 1997;271:587–597
12. Ramsden DA, Baetz K, Wu GE. Conservation of sequence in recombination signal sequence spacers. *Nucleic Acids Res* 1994;22:1785–1796.
13. Nossal BJ, Szenberg A, Ada GL, Austin CM. Single cell studies on 19S antibody production. *J Exp Med* 1964;119:485–501.
14. Wang AC, Wang IYF, McCormick MN, Fudenberg HH. The identity of light chains of monoclonal IgG and monoclonal IgM in one patient. *Immunochemistry* 1969;6:451–459.
15. Nisonoff A, Fudenberg HH, Wilson SK, Hopper JE, Wang AC. Individual antigenic specificiy in immunoglobulins: Relationship to biosynthesis. *Fed Proc* 1972;31:206–209.
16. Gearhart PJ, Sigal NH, Klinman NR. Production of antibodies of identical idiotype but diverse immunoglobulin classes by cells derived from a single stimulated B cell. *Proc Natl Acad Sci USA* 1975;72:1707–1711.
17. Kincade PW, Lawton AR, Bockman DE, Cooper MD. Suppression of immunoglobulin G synthesis as a result of antibody mediated suppression of immunoglobulin M synthesis in chickens. *Proc Natl Acad Sci USA* 1970;67: 1918–1925.
18. Davis MM, Calame K, Early PW, et al. An immunoglobulin heavy-chain gene is formed by at least two recombinational events. *Nature* 1980;283:733–739.
19. Kataoka T, Kawakami T, Takahashi N, Honjo T. Rearrangement of immunoglobulin gamma 1-chain gene and mechanism for heavy-chain class switch. *Proc Natl Acad Sci USA* 1980;77:919–923.
20. Maki R, Traunecker A, Sakano H, Roeder W, Tonegawa S. Exon shuffling generates an immunoglobulin heavy chain gene. *Proc Natl Acad Sci USA* 1980;77: 2138–2142.
21. Honjo T, Kataoka T. Organization of immunoglobulin heavy chain genes and allelic deletion model. *Proc Natl Acad Sci USA* 1978;75:2140–2144.
22. Cory S, Adams JM. Deletions are associated with somatic rearrangement of immunoglobulin heavy chain genes. *Cell* 1980;19:37–51.
23. Tucker PW, Marcu KB, Newell N, Richards J, Blattner FR. Sequence of the cloned gene for the constant region of murine gamma 2b immunoglobulin heavy chain. *Science* 1979;206:1303–1306.
24. Calame K, Rogers J, Early P, et al. Mouse Cmu heavy chain immunoglobulin gene segment contains three intervening sequences separating domains. *Nature* 1980;284:452–455.
25. Tucker PW, Slightom JL, Blattner FR. Mouse IgA heavy chain gene sequence: Implications for evolution of immunoglobulin hinge axons. *Proc Natl Acad Sci USA* 1981;78:7684–7688.
26. Tucker PW, Marcu KB, Slightom JL, Blattner FR. Structure of the constant and 3′ untranslated regions of the murine gamma 2b heavy chain messenger RNA. *Science* 1979;206:1299–1303.
27. Yamawaki KY, Kataoka T, Takahashi N, Obata M, Honjo T. Complete nucleotide sequence of immunoglobulin gamma2b chain gene cloned from newborn mouse DNA. *Nature* 1980;283:786–789.
28. Dard P, Huck S, Frippiat JP, et al. The IGHG3 gene shows a structural polymorphism characterized by different hinge lengths: Sequence of a new 2-exon hinge gene. *Hum Genet* 1997;99:138–141.
29. Yassalli P, Tedghi R, Lisowska-Bernstein B, Tartakoff A, Jaton JC. Evidence fo hydrophobic region within heavy chains of mouse B lymphocyte membrane bound IgM. *Proc Natl Acad Sci USA* 1979;76:5515–5519.
30. Alt FW, Bothwell AL, Knapp M, et al. Synthesis of secreted and membrane bound immunoglobulin mu heavy chains is directed by mRNAs that differ a their 3′ ends. *Cell* 1980;20:293–301.
31. Early P, Rogers J, Davis M, et al. Two mRNAs can be produced from a singl immunoglobulin mu gene by alternative RNA processing pathways. *Cell* 1980 20:313–319.
32. Kemp DJ, Harris AW, Adams JM. Transcripts of the immunoglobulin C mu gen vary in structure and splicing during lymphoid development. *Proc Natl Acad Sc USA* 1980;77:7400–7404.

33. Rogers J, Early P, Carter C, et al. Two mRNAs with different 3' ends encode membrane-bound and secreted forms of immunoglobulin mu chain. *Cell* 1980; 20:303–312.
34. Rogers J, Choi E, Souza L, et al. Gene segments encoding transmembrane carboxyl termini of immunoglobulin gamma chains. *Cell* 1981;26:19–27.
35. Tyler BM, Cowman AF, Gerondakis SD, Adams JM, Bernard O. mRNA for surface immunoglobulin gamma chains encodes a highly conserved transmembrane sequence and a 28-residue intracellular domain. *Proc Natl Acad Sci USA* 1982; 79:2008–2012.
36. Yamawaki KY, Nakai S, Miyata T, Honjo T. Nucleotide sequences of gene segments encoding membrane domains of immunoglobulin gamma chains. *Proc Natl Acad Sci USA* 1982;79:2623–2627.
37. Zhang K, Saxon A, Max EE. Two unusual forms of human immunoglobulin E encoded by alternative RNA splicing of epsilon heavy chain membrane exons. *J Exp Med* 1992;176:233–243.
38. Tsurushita N, Ho L, Korn LJ. Nuclear factors in B lymphoma enhance splicing of mouse membrane-bound mu mRNA in Xenopus oocytes. *Science* 1988;239: 494–497.
39. Galli G, Guise J, Tucker PW, Nevins JR. Poly(A) site choice rather than splice site choice governs the regulated production of IgM heavy-chain RNAs. *Proc Natl Acad Sci USA* 1988;85:2439–2443.
40. Peterson ML, Gimmi ER, Perry RP. The developmentally regulated shift from membrane to secreted mu mRNA production is accompanied by an increase in cleavage-polyadenylation efficiency but no measurable change in splicing efficiency. *Mol Cell Biol* 1991;11:2324–2327.
41. Peterson ML, Perry RP. Regulated production of mu m and mu s mRNA requires linkage of the poly(A) addition sites and is dependent on the length of the mu s-mu m intron. *Proc Natl Acad Sci USA* 1986;83:8883–8887.
42. Tsurushita N, Korn LJ. Effects of intron length on differential processing of mouse mu heavy-chain mRNA. *Mol Cell Biol* 1987;7:2602–2605.
43. Watakabe A, Inoue K, Sâkamoto H, Shimura Y. A secondary structure at the 3' splice site affects the in vitro splicing reaction of mouse immunoglobulin mu chain pre-mRNAs. *Nucleic Acids Res* 1989;17:8159–8169.
44. Lassman CR, Milcarek C. Regulated expression of the mouse gamma 2b Ig H chain gene is influenced by polyA site order and strength. *J Immunol* 1992;148: 2578–2585.
45. Shimizu A, Takahashi N, Yaoita Y, Honjo T. Organization of the constant-region gene family of the mouse immunoglobulin heavy chain. *Cell* 1982;28:499–506.
46. Akahori Y, Kurosawa Y. Nucleotide sequences of all the gamma gene loci of murine immunoglobulin heavy chains. *Genomics* 1997;41:100–104.
47. Kirsch IR, Morton CC, Nakahara K, Leder P. Human immunoglobulin heavy chain genes map to a region of translocations in malignant B lymphocytes. *Science* 1982;216:301–303.
48. Flanagan JG, Rabbitts TH. Arrangement of human immunoglobulin heavy chain constant region genes implies evolutionary duplication of a segment containing gamma, epsilon and alpha genes. *Nature* 1982;300:709–713.
49. Max EE, Battey J, Ney R, Kirsch IR, Leder P. Duplication and deletion in the human immunoglobulin epsilon genes. *Cell* 1982;29:691–699.
50. Battey J, Max EE, McBride WO, Swan D, Leder P. A processed human immunoglobulin epsilon gene has moved to chromosome 9. *Proc Natl Acad Sci USA* 1982;79:5956–5960.
51. Hisajima H, Nishida Y, Nakai S, Takahashi N, Ueda S, Honjo T. Structure of the human immunoglobulin C epsilon 2 gene, a truncated pseudogene: Implications for its evolutionary origin. *Proc Natl Acad Sci USA* 1983;80:2995–2999.
52. Bensmana M, Huck S, Lefranc G, Lefranc MP. The human immunoglobulin pseudo-gamma IGHGP gene shows no major structural defect. *Nucleic Acids Res* 1988;16:3108.
53. Bottaro A, DeMarchi M, Migone N, Carbonara AO. Pulsed-field gel analysis of human immunoglobulin heavy-chain constant region gene deletions reveals the extent of unmapped regions within the locus. *Genomics* 1989;4:505–508.
54. Hofker MH, Walter MA, Cox DW. Complete physical map of the human immunoglobulin heavy chain constant region gene complex. *Proc Natl Acad Sci USA* 1989;86:5567–5571.
55. Lefranc MP, Hammarstrom L, Smith CI, Lefranc G. Gene deletions in the human immunoglobulin heavy chain constant region locus: Molecular and immunological analysis. *Immunodefic Rev* 1991;2:265–281.
56. Knight KL, Becker RS. Isolation of genes encoding bovine IgM, IgG, IgA and IgE chains. *Vet Immunol Immunopathol* 1987;17:17–24.
57. Butler JE, Brown WR. The immunoglobulins and immunoglobulin genes of swine. *Vet Immunol Immunopathol* 1994;43:5–12.
58. Hamers-Casterman C, Atarhouch T, Muyldermans S, et al. Naturally occurring antibodies devoid of light chains. *Nature* 1993;363:446–448.
59. Bruggemann M, Free J, Diamond A, Howard J, Cobbold S, Waldmann H. Immunoglobulin heavy chain locus of the rat: Striking homology to mouse antibody genes. *Proc Natl Acad Sci USA* 1986;83:6075–6079.
60. Dahan A, Reynaud CA, Weill JC. Nucleotide sequence of the constant region of a chicken mu heavy chain immunoglobulin mRNA. *Nucleic Acids Res* 1983;11: 5381–5389.
61. Weill JC, Reynaud CA, Lassila O, Pink JR. Rearrangement of chicken immunoglobulin genes is not an ongoing process in the embryonic bursa of Fabricius. *Proc Natl Acad Sci USA* 1986;83:3336–3340.
62. Parvari R, Avivi A, Lentner F, et al. Chicken immunoglobulin gamma heavy chains: Limited VH gene repertoire, combinatorial diversification by D gene segments and evolution of the heavy chain locus. *EMBO J* 1988;7:739–744.
63. Reynaud CA, Dahan A, Anquez V, Weill JC. Somatic hyperconversion diversifies the single Vh gene of the chicken with a high incidence in the D region. *Cell* 1989;59:171–183.
64. Schrenzel MD, King DP, McKnight ML, Ferrick DA. Characterization of horse (*Equus caballus*) immunoglobulin mu chain-encoding genes. *Immunogenetics* 1997;45:386–393.
65. Kokubu F, Hinds K, Litman R, Shamblott MJ, Litman GW. Extensive families of constant region genes in a phylogenetically primitive vertebrate indicate an additional level of immunoglobulin complexity. *Proc Natl Acad Sci USA* 1987;84: 5868–5872.
66. Amemiya CT, Litman GW. Complete nucleotide sequence of an immunoglobulin heavy-chain gene and analysis of immunoglobulin gene organization in a primitive teleost species. *Proc Natl Acad Sci USA* 1990;87:811–815.
67. Wilson MR, Marcuz A, van GF, et al. The immunoglobulin M heavy chain constant region gene of the channel catfish, *Ictalurus punctatus*: an unusual mRNA splice pattern produces the membrane form of the molecule. *Nucleic Acids Res* 1990;18:5227–5233.
68. Litman GW, Murphy K, Berger L, Litman R, Hinds K, Erickson BW. Complete nucleotide sequences of three VH genes in Caiman, a phylogenetically ancient reptile: Evolutionary diversification in coding segments and variation in the structure and organization of recombination elements. *Proc Natl Acad Sci USA* 1985;82:844–848.
69. Yamawaki KY, Honjo T. Nucleotide sequences of variable region segments of the immunoglobulin heavy chain of *Xenopus laevis*. *Nucleic Acids Res* 1987;15:5888.
70. Charlemagne J. Noninbred axolotls use the same unique heavy chain and a limited number of light chains for their anti-2,4-dinitrophenyl antibody responses. *Eur J Immunol* 1987;17:421–424.
71. Davis MM, Kim SK, Hood LE. DNA sequences mediating class switching in alpha immunoglobulins. *Science* 1980;209:1360–1365.
72. Dunnick W, Rabbitts TH, Milstein C. An immunoglobulin deletion mutant with implications for the heavy-chain switch and RNA splicing. *Nature* 1980;286: 669–675.
73. Kataoka T, Miyata T, Honjo T. Repetitive sequences in class-switch recombination regions of immunoglobulin heavy chain genes. *Cell* 1981;23:357–368.
74. Nikaido T, Nakai S, Honjo T. Switch region of immunoglobulin Cmu gene is composed of simple tandem repetitive sequences. *Nature* 1981;292:845–848.
75. Marcu KB, Banerji J, Penncavage NA, Lang R, Arnheim N. 5' flanking region of immunoglobulin heavy chain constant region genes displays length heterogeneity in germlines of inbred mouse strains. *Cell* 1980;22:187–196.
76. Winter E, Krawinkel U, Radbruch A. Directed Ig class switch recombination in activated murine B cells. *EMBO J* 1987;6:1663–1671.
77. Nikaido T, Yamawaki KY, Honjo T. Nucleotide sequences of switch regions of immunoglobulin C epsilon and C gamma genes and their comparison. *J Biol Chem* 1982;257:7322–7329.
78. Mowatt MR, Dunnick WA. DNA sequence of the murine gamma 1 switch segment reveals novel structural elements. *J Immunol* 1986;136:2674–2683.
79. Ravetch JV, Kirsch IR, Leder P. Evolutionary approach to the question of immunoglobulin heavy chain switching: Evidence from cloned human and mouse genes. *Proc Natl Acad Sci USA* 1980;77:6734–6738.
80. Rabbitts TH, Forster A, Milstein CP. Human immunoglobulin heavy chain genes: Evolutionary comparisons of C mu, C delta and C gamma genes and associated switch sequences. *Nucleic Acids Res* 1981;9:4509–4524.
81. Mills FC, Mitchell MP, Harindranath N, Max EE. Human Ig S gamma regions and their participation in sequential switching to IgE. *J Immunol* 1995;155: 3021–3036.
82. Dunnick W, Hertz GZ, Scappino L, Gritzmacher C. DNA sequences at immunoglobulin switch region recombination sites. *Nucleic Acids Res* 1993;21: 365–372.
83. Iwasato T, Shimizu A, Honjo T, Yamagishi H. Circular DNA is excised by immunoglobulin class switch recombination. *Cell* 1990;62:143–149.
84. Matsuoka M, Yoshida K, Maeda T, Usuda S, Sakano H. Switch circular DNA formed in cytokine-treated mouse splenocytes: Evidence for intramolecular DNA deletion in immunoglobulin class switching. *Cell* 1990;62:135–142.
85. von-Schwedler U, Jack HM, Wabl M. Circular DNA is a product of the immunoglobulin class switch rearrangement. *Nature* 1990;345:452–456.
86. Mills FC, Thyphronitis G, Finkelman FD, Max EE. Ig mu-epsilon isotype switch in IL-4-treated human B lymphoblastoid cells. Evidence for a sequential switch. *J Immunol* 1992;149:1075–1085.
87. Zhang K, Mills FC, Saxon A. Switch circles from IL-4-directed epsilon class switching from human B lymphocytes. Evidence for direct, sequential, and multiple step sequential switch from mu to epsilon Ig heavy chain gene. *J Immunol* 1994;152:3427–3435.
88. Malisan F, Briere F, Bridon JM, et al. Interleukin-10 induces immunoglobulin G isotype switch recombination in human CD40-activated naive B lymphocytes. *J Exp Med* 1996;183:937–947.
89. Chu CC, Paul WE, Max EE. Quantitation of immunoglobulin mu-gamma 1 heavy chain switch region recombination by a digestion-circularization polymerase chain reaction method. *Proc Natl Acad Sci USA* 1992;89:6978–6982.
90. Chu CC, Max EE, Paul WE. DNA rearrangement can account for in vitro switching to IgG1. *J Exp Med* 1993;178:1381–1390.
91. Allen RC, Armitage RJ, Conley ME, et al. CD40 ligand gene defects responsible for X-linked hyper-IgM syndrome. *Science* 1993;259:990–993.
92. Xu J, Foy TM, Laman JD, et al. Mice deficient for the CD40 ligand. *Immunity* 1994;1:423–431.
93. Hodgkin PD, Lee JH, Lyons AB. B cell differentiation and isotype switching is related to division cycle number. *J Exp Med* 1996;184:277–281.

94. Siepmann K, Wohlleben G, Gray D. CD40-mediated regulation of interleukin-4 signaling pathways in B lymphocytes. *Eur J Immunol* 1996;26:1544–1552.
95. Cheng G, Cleary AM, Ye ZS, Hong DI, Lederman S, Baltimore D. Involvement of CRAF1, a relative of TRAF, in CD40 signaling. *Science* 1995;267: 1494–1498.
96. Cogne M, Lansford R, Bottaro A, et al. A class switch control region at the 3' end of the immunoglobulin heavy chain locus. *Cell* 1994;77:737–747.
97. Jung S, Rajewsky K, Radbruch A. Shutdown of class switch recombination by deletion of a switch region control element. *Science* 1993;259:984–987.
98. Bottaro A, Lansford R, Xu L, Zhang J, Rothman P, Alt FW. S region transcription per se promotes basal IgE class switch recombination but additional factors regulate the efficiency of the process. *EMBO J* 1994;13:665–674.
99. Lorenz M, Jung S, Radbruch A. Switch transcripts in immunoglobulin class switching. *Science* 1995;267:1825–1828.
100. Reaban ME, Griffin JA. Induction of RNA-stabilized DNA conformers by transcription of an immunoglobulin switch region. *Nature* 1990;348:342–344.
101. Daniels GA, Lieber MR. RNA:DNA complex formation upon transcription of immunoglobulin switch regions: implications for the mechanism and regulation of class switch recombination. *Nucleic Acids Res* 1995;23:5006–5011.
102. Snapper CM, Marcu KB, Zelazowski P. The immunoglobulin class switch: beyond "accessibility." *Immunity* 1997;6:217–223.
103. Ott DE, Alt FW, Marcu KB. Immunoglobulin heavy chain switch region recombination within a retroviral vector in murine pre-B cells. *EMBO J* 1987;6:577–584.
104. Leung H, Maizels N. Regulation and targeting of recombination in extrachromosomal substrates carrying immunoglobulin switch region sequences. *Mol Cell Biol* 1994;14:1450–1458.
105. Daniels GA, Lieber MR. Strand specificity in the transcriptional targeting of recombination at immunoglobulin switch sequences. *Proc Natl Acad Sci USA* 1995;92:5625–5629.
106. Du J, Zhu Y, Shanmugam A, Kenter AL. Analysis of immunoglobulin S gamma3 recombination breakpoints by PCR: Implications for the mechanism of isotype switching. *Nucleic Acids Res* 1997;25:3066–3074.
107. Williams M, Maizels N. LR1, a lipopolysaccharide-responsive factor with binding sites in the immunoglobulin switch regions and heavy-chain enhancer. *Genes Dev* 1991;5:2353–2361.
108. Williams M, Hanakahi LA, Maizels N. Purification and properties of LR1, an inducible DNA binding protein from mammalian B lymphocytes. *J Biol Chem* 1993;268:13731–13737.
109. Hanakahi LA, Dempsey LA, Li MJ, Maizels N. Nucleolin is one component of the B cell-specific transcription factor and switch region binding protein, LR1. *Proc Natl Acad Sci USA* 1997;94:3605–3610.
110. Mizuta TR, Fukita Y, Miyoshi T, Shimizu A, Honjo T. Isolation of cDNA encoding a binding protein specific to 5'-phosphorylated single-stranded DNA with G-rich sequences. *Nucleic Acids Res* 1993;21:1761–1766.
111. Wuerffel RA, Nathan AT, Kenter AL. Detection of an immunoglobulin switch region-specific DNA-binding protein in mitogen-stimulated mouse splenic B cells. *Mol Cell Biol* 1990;10:1714–1718.
112. Kenter AL, Wuerffel R, Sen R, Jamieson CE, Merkulov GV. Switch recombination breakpoints occur at nonrandom positions in the S gamma tandem repeat. *J Immunol* 1993;151:4718–4731.
113. Ma L, Hu B, Kenter AL. Ig S-gamma-specific DNA binding protein SNAP is related to the helix-loop-helix transcription factor E47. *Int Immunol* 1997;9: 1021–1029.
114. Snapper CM, Zelazowski P, Rosas FR, et al. B cells from p50/NF-kappa B knockout mice have selective defects in proliferation, differentiation, germ-line CH transcription, and Ig class switching. *J Immunol* 1996;156:183–191.
115. Goldfarb AN, Flores JP, Lewandowska K. Involvement of the E2A basic helix-loop-helix protein in immunoglobulin heavy chain class switching. *Mol Immunol* 1996;33:947–956.
116. Lyon CJ, Aguilera RJ. Purification and characterization of the immunoglobulin switch sequence-specific endonuclease (Endo-SR) from bovine spleen. *Mol Immunol* 1997;34:209–219.
116a. Rolink A, Melchers F, Andersson J. The SCID but not the RAG-2 gene product is required for S mu-S epsilon heavy chain class switching. *Immunity* 1996;5: 319–330.
116b. Casellas R, Nussenzweig A, Wuerffel R, et al. Ku80 is required for immunoglobulin isotype switching. *EMBO J* 1998;17:2404–2411.
117. Jessberger R, Wabl M, Borggrefe T. Biochemical studies of class switch recombination. *Curr Top Microbiol Immunol* 1996;217:191–202.
118. Borggrefe T, Wabl M, Akhmedov AT, Jessberger R. A B-Cell specific DNA recombination complex. *J Biol Chem* 1998;in press.
119. Siebenkotten G, Esser C, Wabl M, Radbruch A. The murine IgG1/IgE class switch program. *Eur J Immunol* 1992;22:1827–1834.
120. Mandler R, Finkelman FD, Levine AD, Snapper CM. IL-4 induction of IgE class switching by lipopolysaccharide-activated murine B cells occurs predominantly through sequential switching. *J Immunol* 1993;150:407–418.
121. Jung S, Siebenkotten G, Radbruch A. Frequency of immunoglobulin E class switching is autonomously determined and independent of prior switching to other classes. *J Exp Med* 1994;179:2023–2026.
122. Baskin B, Islam KB, Evengard B, Emtestam L, Smith CI. Direct and sequential switching from mu to epsilon in patients with Schistosoma mansoni infection and atopic dermatitis. *Eur J Immunol* 1997;27:130–135.
123. Greenberg R, Lang RB, Diamond MS, Marcu KB. A switch region inversion contributes to the aberrant rearrangement of a mu immunoglobulin heavy chain gene in MPC-11 cells. *Nucleic Acids Res* 1982;10:7751–7761.
124. Jack HM, McDowell M, Steinberg CM, Wabl M. Looping out and deletion mechanism for the immunoglobulin heavy-chain class switch. *Proc Natl Acad Sci USA* 1988;85:1581–1585.
125. Laffan M, Luzzatto L. Anomalous rearrangements of the immunoglobulin heavy chain genes in human leukemias support the loop-out mechanism of class switch. *J Clin Invest* 1992;90:2299–2303.
126. Knight KL, Kingzette M, Crane MA, Zhai SK. Transchromosomally derived Ig heavy chains. *J Immunol* 1995;155:684–691.
127. Maki R, Roeder W, Traunecker A, et al. The role of DNA rearrangement and alternative RNA processing in the expression of immunoglobulin delta genes. *Cell* 1981;24:353–365.
128. Shimizu A, Nussenzweig MC, Han H, Sanchez M, Honjo T. Trans-splicing as a possible molecular mechanism for the multiple isotype expression of the immunoglobulin gene. *J Exp Med* 1991;173:1385–1393.
129. Shimizu A, Honjo T. Synthesis and regulation of trans-mRNA encoding the immunoglobulin epsilon heavy chain. *FASEB J* 1993;7:149–154.
130. Mizuta TR, Suzuki N, Shimizu A, Honjo T. Duplicated variable region genes account for double isotype expression in a human leukemic B-cell line that gives rise to single isotype-expressing cells. *J Biol Chem* 1991;266:12514–12521.
131. Chen YW, Word C, Dev V, Uhr JW, Vitetta ES, Tucker PW. Double isotype production by a neoplastic B cell line. II. Allelically excluded production of mu and gamma 1 heavy chains without CH gene rearrangement. *J Exp Med* 1986;164: 562–579.
132. Akahori Y, Kurosawa Y, Kamachi Y, Torii S, Matsuoka H. Presence of immunoglobulin (Ig) M and IgG double isotype-bearing cells and defect of switch recombination in hyper IgM immunodeficiency. *J Clin Invest* 1990;85: 1722–1727.
133. Kunimoto DY, Sneller MC, Claflin L, Mushinski JF, Strober W. Molecular analysis of double isotype expression in IgA switching. *J Immunol* 1993;150: 1338–1347.
134. Hieter PA, Max EE, Seidman JG, Maizel JJ, Leder P. Cloned human and mouse kappa immunoglobulin constant and J region genes conserve homology in functional segments. *Cell* 1980;22:197–207.
135. Hieter PA, Maizel JJ, Leder P. Evolution of human immunoglobulin kappa J region genes. *J Biol Chem* 1982;257:1516–1522.
136. Hieter PA, Korsmeyer SJ, Waldmann TA, Leder P. Human immunoglobulin kappa light-chain genes are deleted or rearranged in lambda-producing B cells. *Nature* 1981;290:368–372.
137. Durdik J, Moore MW, Selsing E. Novel kappa light-chain gene rearrangements in mouse lambda light chain-producing B lymphocytes. *Nature* 1984;307: 749–752.
138. Siminovitch KA, Bakhshi A, Goldman P, Korsmeyer SJ. A uniform deleting element mediates the loss of kappa genes in human B cells. *Nature* 1985;316: 260–262.
139. Klobeck HG, Zachau HG. The human CK gene segment and the kappa deleting element are closely linked. *Nucleic Acids Res* 1986;14:4591–4603.
140. Muller B, Stappert H, Reth M. A physical map and analysis of the murine C kappa-RS region show the presence of a conserved element. *Eur J Immunol* 1990;20:1409–1411.
141. Tonegawa S, Brack C, Hozumi N, Schuller R. Cloning of an immunoglobulin variable region gene from mouse embryo. *Proc Natl Acad Sci USA* 1977;74: 3518–3522.
142. Miller J, Ogden S, McMullen M, Andres H, Storb U. The order and orientation of mouse lambda-genes explain lambda-rearrangement patterns. *J Immunol* 1988;141:2497–2502.
143. Storb U, Haasch D, Arp B, Sanchez P, Cazenave PA, Miller J. Physical linkage of mouse lambda genes by pulsed-field gel electrophoresis suggests that the rearrangement process favors proximate target sequences. *Mol Cell Biol* 1989;9: 711–718.
144. Sanchez P, Marche PN, Rueff JD, Cazenave PA. Mouse V lambda x gene sequence generates no junctional diversity and is conserved in mammalian species. *J Immunol* 1990;144:2816–2820.
145. Miller J, Selsing E, Storb U. Structural alterations in J regions of mouse immunoglobulin lambda genes are associated with differential gene expression. *Nature* 1982;295:428–430.
146. Scott CL, Mushinski JF, Huppi K, Weigert M, Potter M. Amplification of immunoglobulin lambda constant genes in populations of wild mice. *Nature* 1982;300:757–760.
147. Hieter PA, Hollis GF, Korsmeyer SJ, Waldmann TA, Leder P. Clustered arrangement of immunoglobulin lambda constant region genes in man. *Nature* 1981; 294:536–540.
148. Vasicek TJ, Leder P. Structure and expression of the human immunoglobulin lambda genes. *J Exp Med* 1990;172:609–620.
149. Kawasaki K, Minoshima S, Nakato E, et al. One-megabase sequence analysis of the human immunoglobulin lambda gene locus. *Genome Res* 1997;7:250–261.
150. Dariavach P, Lefranc G, Lefranc MP. Human immunoglobulin C lambda 6 gene encodes the Kern+Oz-lambda chain and C lambda 4 and C lambda 5 are pseudogenes. *Proc Natl Acad Sci USA* 1987;84:9074–9078.
151. Stiernholm NB, Verkoczy LK, Berinstein NL. Rearrangement and expression of the human psi C lambda 6 gene segment results in a surface Ig receptor with a truncated light chain constant region. *J Immunol* 1995;154:4583–4591.
152. Taub RA, Hollis GF, Hieter PA, Korsmeyer S, Waldmann TA, Leder P. Variable amplification of immunoglobulin lambda light-chain genes in human populations. *Nature* 1983;304:172–174.
153. Hollis GF, Hieter PA, McBride OW, Swan D, Leder P. Processed genes: A dis-

persed human immunoglobulin gene bearing evidence of RNA-type processing. *Nature* 1982;296:321–325.
154. Sakaguchi N, Melchers F. Lambda 5, a new light-chain-related locus selectively expressed in pre-B lymphocytes. *Nature* 1986;324:579–582.
155. Kudo A, Sakaguchi N, Melchers F. Organization of the murine Ig-related lambda 5 gene transcribed selectively in pre-B lymphocytes. *EMBO J* 1987;6:103–107.
156. Kudo A, Melchers F. A second gene, VpreB in the lambda 5 locus of the mouse, which appears to be selectively expressed in pre-B lymphocytes. *EMBO J* 1987;6:2267–2272.
157. Sakaguchi N, Berger CN, Melchers F. Isolation of a cDNA copy of an RNA species expressed in murine pre-B cells. *EMBO J* 1986;5:2139–2147.
158. Dul JL, Argon Y, Winkler T, ten Boekel E, Melchers F, Martensson IL. The murine VpreB1 and VpreB2 genes both encode a protein of the surrogate light chain and are co-expressed during B cell development. *Eur J Immunol* 1996;26:906–913.
159. Hagiwara S, Tsunetsugu YY, Kimoto H, Takemori T. Expression of Vpre-B3 (8HS-20) molecules by alternative RNA processing. *Int Immunol* 1996;8: 1237–1244.
160. Tsubata T, Reth M. The products of pre-B cell-specific genes (lambda 5 and VpreB) and the immunoglobulin mu chain form a complex that is transported onto the cell. *J Exp Med* 1990;172:973–976.
161. Cherayil BJ, Pillai S. The omega/lambda 5 surrogate immunoglobulin light chain is expressed on the surface of transitional B lymphocytes in murine bone marrow. *J Exp Med* 1991;173:111–116.
162. Pillai S, Baltimore D. The omega and iota surrogate immunoglobulin light chains. *Curr Top Microbiol Immunol* 1988;137:136–139.
163. Tsubata T, Tsubata R, Reth M. Crosslinking of the cell surface immunoglobulin (mu-surrogate light chains complex) on pre-B cells induces activation of V gene rearrangements at the immunoglobulin kappa locus. *Int Immunol* 1992;4: 637–641.
164. Bauer TJ, McDermid HE; Budarf ML, Van KM, Blomberg BB. Physical location of the human immunoglobulin lambda-like genes, 14.1, 16.1, and 16.2. *Immunogenetics* 1993;38:387–399.
165. Bossy D, Milili M, Zucman J, Thomas G, Fougereau M, Schiff C. Organization and expression of the lambda-like genes that contribute to the mu-psi light chain complex in human pre-B cells. *Int Immunol* 1991;3:1081–1090.
166. Evans RJ, Hollis GF. Genomic structure of the human Ig lambda 1 gene suggests that it may be expressed as an Ig lambda 14.1-like protein or as a canonical B cell Ig lambda light chain: Implications for Ig lambda gene evolution. *J Exp Med* 1991;173:305–311.
167. Schiff C, Milili M, Bossy D, Fougereau M. Organization and expression of the pseudo-light chain genes in human B-cell ontogeny. *Int Rev Immunol* 1992;8: 135–145.
168. Frippiat JP, Williams SC, Tomlinson IM, et al. Organization of the human immunoglobulin lambda light-chain locus on chromosome 22q11.2. *Hum Mol Genet* 1995;4:983–991.
169. Van Ness BG, Coleclough C, Perry RP, Weigert M. DNA between variable and joining gene segments of immunoglobulin kappa light chain is frequently retained in cells that rearrange the kappa locus. *Proc Natl Acad Sci USA* 1982; 79:262–266.
170. Hochtl J, Müller CR, Zachau HG. Recombined flanks of the variable and joining segments of immunoglobulin genes. *Proc Natl Acad Sci USA* 1982;79: 1383–1387.
171. Hochtl J, Zachau HG. A novel type of aberrant recombination in immunoglobulin genes and its implications for V-J joining mechanism. *Nature* 1983;302: 260–263.
172. Feddersen RM, Van NB. Double recombination of a single immunoglobulin kappa-chain allele: Implications for the mechanism of rearrangement. *Proc Natl Acad Sci USA* 1985;82:4793–4797.
173. Lewis S, Gifford A, Baltimore D. Joining of V kappa to J kappa gene segments in a retroviral vector introduced into lymphoid cells. *Nature* 1984;308:425–428.
174. Hesse JE, Lieber MR, Gellert M, Mizuuchi K. Extrachromosomal DNA substrates in pre-B cells undergo inversion or deletion at immunoglobulin V-(D)-J joining signals. *Cell* 1987;49:775–783.
175. Shapiro MA, Weigert M. How immunoglobulin V kappa genes rearrange. *J Immunol* 1987;139:3834–3839.
176. Shimizu T, Iwasato T, Yamagishi H. Deletions of immunoglobulin C kappa region characterized by the circular excision products in mouse splenocytes. *J Exp Med* 1991;173:1065–1072.
177. Selsing E, Voss J, Storb U. Immunoglobulin gene "remnant" DNA—implications for antibody gene recombination. *Nucleic Acids Res* 1984;12:4229–4246.
178. Maeda T, Sugiyama H, Tani Y, et al. Start of mu-chain production by the further two-step rearrangements of immunoglobulin heavy chain genes on one chromosome from a DJH/DJH configuration in an Abelson virus-transformed cell line: Evidence of secondary DJH complex formation. *J Immunol* 1987;138:2305–2310.
179. Reth M, Gehrmann P, Petrac E, Wiese P. A novel VH to VHDJH joining mechanism in heavy-chain-negative (null) pre-B cells results in heavy-chain production. *Nature* 1986;322:840–842.
180. Covey LR, Ferrier P, Alt FW. VH to VHDJH rearrangement is mediated by the internal VH heptamer. *Int Immunol* 1990;2:579–583.
181. Chen C, Nagy Z, Prak EL, Weigert M. Immunoglobulin heavy chain gene replacement: A mechanism of receptor editing. *Immunity* 1995;3:747–755.
182. Schuler W, Weiler IJ, Schuler A, et al. Rearrangement of antigen receptor genes is defective in mice with severe combined immune deficiency. *Cell* 1986;46: 963–972.
183. Lafaille JJ, DeCloux A, Bonneville M, Takagaki Y, Tonegawa S. Junctional sequences of T cell receptor gamma delta genes: implications for gamma delta T cell lineages and for a novel intermediate of V-(D)-J joining. *Cell* 1989;59: 859–870.
184. McCormack WT, Tjoelker LW, Carlson LM, et al. Chicken IgL gene rearrangement involves deletion of a circular episome and addition of single nonrandom nucleotides to both coding segments. *Cell* 1989;56:785–791.
185. Schuler W, Ruetsch NR, Amsler M, Bosma MJ. Coding joint formation of endogenous T cell receptor genes in lymphoid cells from scid mice: Unusual P-nucleotide additions in VJ-coding joints. *Eur J Immunol* 1991;21:589–596.
186. Alt FW, Baltimore D. Joining of immunoglobulin heavy chain gene segments: Implications from a chromosome with evidence of three D-JH fusions. *Proc Natl Acad Sci USA* 1982;79:4118–4122.
187. Alt F, Rosenberg N, Lewis S, Thomas E, Baltimore D. Organization and reorganization of immunoglobulin genes in A-MULV-transformed cells: Rearrangement of heavy but not light chain genes. *Cell* 1981;27:391–390.
188. Lewis S, Rosenberg N, Alt F, Baltimore D. Continuing kappa-gene rearrangement in a cell line transformed by Abelson murine leukemia virus. *Cell* 1982;30: 807–816.
189. Reth MG, Ammirati P, Jackson S, Alt FW. Regulated progression of a cultured pre-B-cell line to the B-cell stage. *Nature* 1985;317:353–355.
190. Lewis S, Gifford A, Baltimore D. DNA elements are asymmetrically joined during the site-specific recombination of kappa immunoglobulin genes. *Science* 1985;228:677–685.
191. Lieber MR, Hesse JE, Mizuuchi K, Gellert M. Developmental stage specificity of the lymphoid V(D)J recombination activity. *Genes Dev* 1987;1:751–761.
192. Lieber MR, Hesse JE, Lewis S, et al. The defect in murine severe combined immune deficiency: Joining of signal sequences but not coding segments in V(D)J recombination. *Cell* 1988;55:7–16.
193. Lewis SM, Hesse JE, Mizuuchi K, Gellert M. Novel strand exchanges in V(D)J recombination. *Cell* 1988;55:1099–1107.
194. Akira S, Okazaki K, Sakano H. Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining. *Science* 1987;238:1134–1138.
195. Hesse JE, Lieber MR, Mizuuchi K, Gellert M. V(D)J recombination: a functional definition of the joining signals. *Genes Dev* 1989;3:1053–1061.
196. Roth DB, Menetski JP, Nakajima PB, Bosma MJ, Gellert M. V(D)J recombination: Broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. *Cell* 1992;70:983–991.
197. Roth DB, Nakajima PB, Menetski JP, Bosma MJ, Gellert M. V(D)J recombination in mouse thymocytes: Double-strand breaks near T cell receptor delta rearrangement signals. *Cell* 1992;69:41–53.
198. Roth DB, Zhu C, Gellert M. Characterization of broken DNA molecules associated with V(D)J recombination. *Proc Natl Acad Sci USA* 1993;90:10788–10792.
199. Schlissel M, Constantinescu A, Morrow T, Baxter M, Peng A. Double-strand signal sequence breaks in V(D)J recombination are blunt, 5'-phosphorylated, RAG-dependent, and cell cycle regulated. *Genes Dev* 1993;7:2520–2532.
200. Zhu C, Roth DB. Characterization of coding ends in thymocytes of scid mice: Implications for the mechanism of V(D)J recombination. *Immunity* 1995;2: 101–112.
201. Ramsden DA, Gellert M. Formation and resolution of double-strand break intermediates in V(D)J rearrangement. *Genes Dev* 1995;9:2409–2420.
202. Lewis SM. P nucleotide insertions and the resolution of hairpin DNA structures in mammalian cells. *Proc Natl Acad Sci USA* 1994;91:1332–1336.
203. Desiderio S, Baltimore D. Double-stranded cleavage by cell extracts near recombinational signal sequences of immunoglobulin genes. *Nature* 1984;308: 860–862.
204. Kataoka T, Kondo S, Nishi M, Kodaira M, Honjo T. Isolation and characterization of endonuclease J: A sequence-specific endonuclease cleaving immunoglobulin genes. *Nucleic Acids Res* 1984;12:5995–6010.
205. Hope TJ, Aguilera RJ, Minie ME, Sakano H. Endonucleolytic activity that cleaves immunoglobulin recombination sequences. *Science* 1986;231:1141–1145.
206. Schatz DG, Baltimore D. Stable expression of immunoglobulin gene V(D)J recombinase activity by gene transfer into 3T3 fibroblasts. *Cell* 1988;53: 107–115.
207. Oettinger MA, Schatz DG, Gorka C, Baltimore D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 1990;248: 1517–1523.
208. Bernstein RM, Schluter SF, Bernstein H, Marchalonis JJ. Primordial emergence of the recombination activating gene 1 (RAG1): Sequence of the complete shark gene indicates homology to microbial integrases. *Proc Natl Acad Sci USA* 1996; 93:9454–9459.
209. Mombaerts P, Iacomini J, Johnson RS, et al. RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* 1992;68:869–877.
210. Shinkai Y, Rathbun G, Lam KP, et al. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* 1992;68:855–867.
211. Schwarz K, Gauss GH, Ludwig L, et al. RAG mutations in human B cell-negative SCID. *Science* 1996;274:97–99.
212. Sadofsky MJ, Hesse JE, McBlane JF, Gellert M. Expression and V(D)J recombination activity of mutated RAG-1 proteins. *Nucleic Acids Res* 1993;21: 5644–5650.
213. McBlane JF, van GD, Ramsden DA, et al. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell* 1995;83:387–395.
214. van Gent DC, Mizuuchi K, Gellert M. Similarities between initiation of V(D)J recombination and retroviral integration. *Science* 1996;271:1592–1594.

215. van Gent DC, Ramsden DA, Gellert M. The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination. *Cell* 1996;85:107–113.
216. Hiom K, Gellert M. A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. *Cell* 1997;88:65–72.
217. Sawchuk DJ, Weis GF, Malik S, et al. V(D)J recombination: Modulation of RAG1 and RAG2 cleavage activity on 12/23 substrates by whole cell extract and DNA-bending proteins. *J Exp Med* 1997;185:2025–2032.
218. Cuomo CA, Mundy CL, Oettinger MA. DNA sequence and structure requirements for cleavage of V(D)J recombination signal sequences. *Mol Cell Biol* 1996;16:5683–5690.
219. Ramsden DA, McBlane JF, van Gent DC, Gellert M. Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. *EMBO J* 1996;15:3197–3206.
220. Cortes P, Weis GF, Misulovin Z, et al. In vitro V(D)J recombination: Signal joint formation. *Proc Natl Acad Sci USA* 1996;93:14008–14013.
221. Leu TMJ, Eastman QM, Schatz DG. Coding joint formation in a cell-free V(D)J recombination system. *Immunity* 1997;7:303–314.
222. Ramsden DA, Paull TT, Gellert M. Cell-free V(D)J recombination. *Nature* 1997;388:488–491.
223. Agrawal A, Schatz DG. RAG1 and RAG2 form a stable postcleavage synaptic complex with DNA containing signal ends in V(D)J recombination. *Cell* 1997; 89:43–53.
224. Lin WC, Desiderio S. Regulation of V(D)J recombination activator protein RAG-2 by phosphorylation. *Science* 1993;260:953–959.
225. Lin WC, Desiderio S. Cell cycle regulation of V(D)J recombination-activating protein RAG-2. *Proc Natl Acad Sci USA* 1994;91:2733–2737.
226. Li Z, Dordai DI, Lee J, Desiderio S. A conserved degradation signal regulates RAG-2 accumulation during cell division and links V(D)J recombination to the cell cycle. *Immunity* 1996;5:575–589.
227. Difilippantonio MJ, McMahan CJ, Eastman QM, Spanopoulou E, Schatz DG. RAG1 mediates signal sequence recognition and recruitment of RAG2 in V(D)J recombination. *Cell* 1996;87:253–262.
228. Spanopoulou E, Zaitseva F, Wang FH, Santagata S, Baltimore D, Panayotou G. The homeodomain region of Rag-1 reveals the parallel mechanisms of bacterial and V(D)J recombination. *Cell* 1996;87:263–276.
229. Chun JJ, Schatz DG, Oettinger MA, Jaenisch R, Baltimore D. The recombination activating gene-1 (RAG-1) transcript is present in the murine central nervous system. *Cell* 1991;64:189–200.
230. Carlson LM, Oettinger MA, Schatz DG, et al. Selective expression of RAG-2 in chicken B cells undergoing immunoglobulin gene conversion. *Cell* 1991;64: 201–208.
231. Takeda S, Masteller EL, Thompson CB, Buerstedde JM. RAG-2 expression is not essential for chicken immunoglobulin gene conversion. *Proc Natl Acad Sci USA* 1992;89:4023–4027.
232. Chen J, Lansford R, Stewart V, Young F, Alt FW. RAG-2-deficient blastocyst complementation: An assay of gene function in lymphocyte development. *Proc Natl Acad Sci USA* 1993;90:4528–4532.
233. Taccioli GE, Rathbun G, Oltz E, Stamato T, Jeggo PA, Alt FW. Impairment of V(D)J recombination in double-strand break repair mutants. *Science* 1993;260: 207–210.
234. Troelstra C, Jaspers NGJ. Ku starts at the end. *Curr Biol* 1994;4:1149–1151.
234a. Wu X, Lieber MR. Protein-protein and protein-DNA interaction regions within the DNA end-binding protein Ku70-Ku86. *Mol Cell Biol* 1996;16:5186–5193.
234b. Yaneva M, Kowalewski T, Lieber MR. Interaction of DNA-dependent protein kinase with DNA and with Ku: Biochemical and atomic-force microscopy studies. *EMBO J* 1997;16:5098–5112.
235. Taccioli GE, Gottlieb TM, Blunt T, et al. Ku80: Product of the XRCC5 gene and its role in DNA repair and V(D)J recombination. *Science* 1994;265:1442–1445.
236. Errami A, Smider V, Rathmell WK, et al. Ku86 defines the genetic defect and restores X-ray resistance and V(D)J recombination to complementation group 5 hamster cell mutants. *Mol Cell Biol* 1996;16:1519–1526.
237. Blunt T, Finnie NJ, Taccioli GE, et al. Defective DNA-dependent protein kinase activity is linked to V(D)J recombination and DNA repair defects associated with the murine scid mutation. *Cell* 1995;80:813–823.
238. Shin EK, Perryman LE, Meek K. A kinase-negative mutation of DNA-PK(CS) in equine SCID results in defective coding and signal joint formation. *J Immunol* 1997;158:3565–3569.
239. Gu Y, Jin S, Gao Y, Weaver DT, Alt FW. Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V(D)J recombination. *Proc Natl Acad Sci USA* 1997;94:8076–8081.
240. de Vries E, van DW, Bergsma WG, et al. HeLa nuclear protein recognizing DNA termini and translocating on DNA forming a regular DNA-multimeric protein complex. *J Mol Biol* 1989;208:65–78.
241. Tuteja N, Tuteja R, Ochem A, et al. Human DNA helicase II: A novel DNA unwinding enzyme identified as the Ku autoantigen. *EMBO J* 1994;13: 4991–5001.
242. Li Z, Otevrel T, Gao Y, et al. The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. *Cell* 1995;83: 1079–1089.
243. Critchlow SE, Bowater RP, Jackson SP. Mammalian DNA double-strand break repair protein XRCC4 interacts with DNA ligase IV. *Curr Biol* 1997;7:588–598.
244. Grawunder U, Wilm M, Wu XK, P., Wilson TE, Mann M, Lieber MR. Activity of DNA ligase IV stimulate by complex formation with XRCC4 in mammalian cells. *Nature* 1997;388:492.
244a. Leber R, Wise TW, Mizuta R, Meek K. The XRCC4 gene product is a target for and interacts with the DNA-dependent protein kinase. *J Biol Chem* 1998;273: 1794-1801.
245. Hsieh JJ, Henkel T, Salmon P, Robey E, Peterson MG, Hayward SD. Truncated mammalian Notch1 activates CBF1/RBPJk-repressed genes by a mechanism resembling that of Epstein-Barr virus EBNA2. *Mol Cell Biol* 1996;16:952–959.
246. Aguilera RJ, Akira S, Okazaki K, Sakano H. A pre-B cell nuclear protein that specifically interacts with the immunoglobulin V-J recombination sequences. *Cell* 1987;51:909–917.
247. Miyake S, Sugiyama H, Tani Y, Fukuda T, Kishimoto S. Identification of a recombinational signal sequence-specific DNA-binding protein(s) of Mr 115,000 in the nuclear extracts from immature lymphoid cell lines. *J Immunogenet* 1990;17:67–75.
248. Shirakata M, Huppi K, Usuda S, Okazaki K, Yoshida K, Sakano H. HMG1-related DNA-binding protein isolated with V-(D)-J recombination signal probes. *Mol Cell Biol* 1991;11:4528–4536.
249. Li M, Morzycka WE, Desiderio SV. NBP, a protein that specifically binds an enhancer of immunoglobulin gene rearrangement: Purification and characterization. *Genes Dev* 1989;3:1801–1813.
250. Halligan BD, Teng M, Guilliams TG, Nauert JB, Halligan NL. Cloning of the murine cDNA encoding VDJP, a protein homologous to the large subunit of replication factor C and bacterial DNA ligases. *Gene* 1995;161:217–222.
251. Guilliams TG, Teng M, Halligan BD. Site directed DNA joining. *Biochimie* 1997;79:13–22.
252. Wu LC, Mak CH, Dear N, Boehm T, Foroni L, Rabbitts TH. Molecular cloning of a zinc finger protein which binds to the heptamer of the signal sequence for V(D)J recombination. *Nucleic Acids Res* 1993;21:5067–5073.
253. Mak CH, Strandtmann J, Wu LC. The V(D)J recombination signal sequence and kappa B binding protein Rc binds DNA as dimers and forms multimeric structures with its DNA ligands. *Nucleic Acids Res* 1994;22:383–390.
254. Muegge K, West M, Durum SK. Recombination sequence-binding protein in thymocytes undergoing T-cell receptor gene rearrangement. *Proc Natl Acad Sci USA* 1993;90:4151–4155.
255. Jarvis CD, Geiman T, Vila SM, et al. A novel putative helicase produced in early murine lymphocytes. *Gene* 1996;169:203–207.
256. Kenter AL, Tredup J. High expression of a 3′–5′ exonuclease activity is specific to B lymphocytes. *Mol Cell Biol* 1991;11:4398–4404.
257. Feeney AJ. Lack of N regions in fetal and neonatal mouse immunoglobulin V-D-J junctional sequences. *J Exp Med* 1990;172:1377–1390.
258. Victor KD, Capra JD. An apparently common mechanism of generating antibody diversity: Length variation of the VL-JL junction. *Mol Immunol* 1994;31:39–46.
259. Gilfillan S, Dierich A, Lemeur M, Benoist C, Mathis D. Mice lacking TdT: Mature animals with an immature lymphocyte repertoire. *Science* 1993;261: 1175–1178.
260. Komori T, Okada A, Stewart V, Alt FW. Lack of N regions in antigen receptor variable region genes of TdT- deficient lymphocytes. *Science* 1993;261:1171–1175.
261. Landau NR, Schatz DG, Rosa M, Baltimore D. Increased frequency of N-region insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase retroviral expression vector. *Mol Cell Biol* 1987;7:3237–3243.
262. Kallenbach S, Doyen N, Fanton dAM, Rougeon F. Three lymphoid-specific factors account for all junctional diversity characteristic of somatic assembly of T-cell receptor and immunoglobulin genes. *Proc Natl Acad Sci USA* 1992; 89:2799–2803.
263. Bentolila LA, Wu GE, Nourrit F, Fanton dAM, Rougeon F, Doyen N. Constitutive expression of terminal deoxynucleotidyl transferase in transgenic mice is sufficient for N region diversity to occur at any Ig locus throughout B cell differentiation. *J Immunol* 1997;158:715–723.
264. Hiramatsu R, Akagi K, Matsuoka M, et al. The 3′ enhancer region determines the B/T specificity and pro-B/pre-B specificity of immunoglobulin V kappa-J kappa joining. *Cell* 1995;83:1113–1123.
265. Wasserman R, Li YS, Hardy RR. Down-regulation of terminal deoxynucleotidyl transferase by Ig heavy chain in B lineage cells. *J Immunol* 1997;158:1133–1138.
266. Gilfillan S, Bachmann M, Trembleau S, et al. Efficient immune responses in mice lacking N-region diversity. *Eur J Immunol* 1995;25:3115–3122.
267. Gavin MA, Bevan MJ. Increased peptide promiscuity provides a rationale for the lack of N regions in the neonatal T cell repertoire. *Immunity* 1995;3:793–800.
268. Rolink A, Melchers F. B lymphopoiesis in the mouse. *Adv Immunol* 1993;53: 123–156.
269. Hardy RR, Hayakawa K. B-lineage differentiation stages resolved by multiparameter flow cytometry. *Ann N Y Acad Sci* 1995;764:19–24.
270. Melchers F, Rolink A, Grawunder U, et al. Positive and negative selection events during B lymphopoiesis. *Curr Opin Immunol* 1995;7:214–227.
271. Li YS, Wasserman R, Hayakawa K, Hardy RR. Identification of the earliest B lineage stage in mouse bone marrow. *Immunity* 1996;5:527–535.
272. Karasuyama H, Rolink A, Melchers F. A complex of glycoproteins is associated with VpreB/lambda 5 surrogate light chain on the surface of mu heavy chain-negative early precursor B cell lines. *J Exp Med* 1993;178:469–478.
273. Reth MG, Alt FW. Novel immunoglobulin heavy chains are produced from DJH gene segment rearrangements in lymphoid cells. *Nature* 1984;312:418–423.
274. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. *Cell* 1985;40:271–281.
275. Mather EL, Perry RP. Transcriptional regulation of immunoglobulin V genes. *Nucleic Acids Res* 1981;9:6855–6867.
276. Takeda S, Zou YR, Bluethmann H, Kitamura D, Muller U, Rajewsky K. Deletion

of the immunoglobulin kappa chain intron enhancer abolishes kappa chain gene rearrangement in cis but not lambda chain gene rearrangement in trans. *EMBO J* 1993;12:2329–2336.
277. Xu Y, Davidson L, Alt FW, Baltimore D. Deletion of the Ig kappa light chain intronic enhancer/matrix attachment region impairs but does not abolish V kappa J kappa rearrangement. *Immunity* 1996;4:377–385.
278. Gorman JR, van der Stoep N, Monroe R, et al. The Ig(kappa) enhancer influences the ratio of Ig(kappa) versus Ig(lambda) B lymphocytes. *Immunity* 1996;5:241–252.
279. Lauster R, Reynaud CA, Martensson IL, et al. Promoter, enhancer and silencer elements regulate rearrangement of an immunoglobulin transgene. *EMBO J* 1993;12:4615–4623.
280. Oltz EM, Alt FW, Lin WC, et al. A V(D)J recombinase-inducible B-cell line: Role of transcriptional enhancer elements in directing V(D)J recombination. *Mol Cell Biol* 1993;13:6223–6230.
281. Kallenbach S, Babinet C, Pournin S, Cavelier P, Goodhardt M, Rougeon F. The intronic immunoglobulin kappa gene enhancer acts independently on rearrangement and on transcription. *Eur J Immunol* 1993;23:1917–1921.
282. Ferradini L, Gu H, De SA, Rajewsky K, Reynaud CA, Weill JC. Rearrangement-enhancing element upstream of the mouse immunoglobulin kappa chain J cluster. *Science* 1996;271:1416–1420.
283. Coleclough C, Perry RP, Karjalainen K, Weigert M. Aberrant rearrangements contribute significantly to the allelic exclusion of immunoglobulin gene expression. *Nature* 1981;290:372–378.
284. Alt FW, Enea V, Bothwell AL, Baltimore D. Activity of multiple light chain genes in murine myeloma cells producing a single, functional light chain. *Cell* 1980; 21:1–12.
285. ten Boekel E, Melchers F, Rolink A. The status of Ig loci rearrangements in single cells from different stages of B cell development. *Int Immunol* 1995;7: 1013–1019.
286. Nottenburg C, St. John T, Weissman IL. Unusual immunoglobulin DNA sequences from the nonexpressed chromosome of mouse normal B lymphocytes: Implications for allelic exclusion and the DNA rearrangement process. *J Immunol* 1987;139:1718–1726.
287. Weaver D, Costantini F, Imanishi KT, Baltimore D. A transgenic immunoglobulin mu gene prevents rearrangement of endogenous genes. *Cell* 1985;42: 117–127.
288. Storb U, Pinkert C, Arp B, et al. Transgenic mice with mu and kappa genes encoding antiphosphorylcholine antibodies. *J Exp Med* 1986;164:627–641.
289. Nussenzweig MC, Shaw AC, Sinn E, et al. Allelic exclusion in transgenic mice that express the membrane form of immunoglobulin mu. *Science* 1987;236: 816–819.
290. Picarella D, Serunian LA, Rosenberg N. Allelic exclusion of membrane but not secreted immunoglobulin in a mature B cell line. *Eur J Immunol* 1991;21:55–62.
291. Kitamura D, Rajewsky K. Targeted disruption of mu chain membrane exon causes loss of heavy-chain allelic exclusion. *Nature* 1992;356:154–156.
292. Loffert D, Ehlich A, Muller W, Rajewsky K. Surrogate light chain expression is required to establish immunoglobulin heavy chain allelic exclusion during early B cell development. *Immunity* 1996;4:133–144.
293. Papavasiliou F, Jankovic M, Suh H, Nussenzweig MC. The cytoplasmic domains of immunoglobulin (Ig) alpha and Ig beta can independently induce the precursor B cell transition and allelic exclusion. *J Exp Med* 1995;182:1389–1394.
294. Papavasiliou F, Misulovin Z, Suh H, Nussenzweig MC. The role of Ig beta in precursor B cell transition and allelic exclusion. *Science* 1995;268:408–411.
295. Grawunder U, Leu TM, Schatz DG, et al. Down-regulation of RAG1 and RAG2 gene expression in preB cells after functional immunoglobulin heavy chain rearrangement. *Immunity* 1995;3:601–608.
296. Schlissel MS, Morrow T. Ig heavy chain protein controls B cell development by regulating germ-line transcription and retargeting V(D)J recombination. *J Immunol* 1994;153:1645–1657.
297. Constantinescu A, Schlissel MS. Changes in locus-specific V(D)J recombinase activity induced by immunoglobulin gene products during B cell development. *J Exp Med* 1997;185:609–620.
298. Shinkai Y, Rathbun G, Lam KP, et al. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* 1992;68: 855–867.
299. Ehlich A, Schaal S, Gu H, Kitamura D, Muller W, Rajewsky K. Immunoglobulin heavy and light chain genes rearrange independently at early stages of B cell development. *Cell* 1993;72:695–704.
300. Spanopoulou E, Roman CA, Corcoran LM, et al. Functional immunoglobulin transgenes guide ordered B-cell differentiation in Rag-1-deficient mice. *Genes Dev* 1994;8:1030–1042.
301. Young F, Ardman B, Shinkai Y, et al. Influence of immunoglobulin heavy- and light-chain expression on B-cell differentiation. *Genes Dev* 1994;8:1043–1057.
301a. Minegishi Y, Coustan SE, Wang YH, Cooper MD, Campana D, Conley ME. Mutations in the human lambda5/14.1 gene result in B cell deficiency and agammaglobulinemia. *J Exp Med* 1998;187:71–77.
302. Pelanda R, Schaal S, Torres RM, Rajewsky K. A prematurely expressed Ig(kappa) transgene, but not V(kappa)J(kappa) gene segment targeted into the Ig(kappa) locus, can rescue B cell development in lambda5-deficient mice. *Immunity* 1996;5:229–239.
303. Reth M, Petrac E, Wiese P, Lobel L, Alt FW. Activation of V kappa gene rearrangement in pre-B cells follows the expression of membrane-bound immunoglobulin heavy chains. *EMBO J* 1987;6:3299–3305.
304. Iglesias A, Lamers M, Kohler G. Expression of immunoglobulin delta chain causes allelic exclusion in transgenic mice. *Nature* 1987;330:482–484.
305. Schlissel MS, Baltimore D. Activation of immunoglobulin kappa gene rearrangement correlates with induction of germline kappa gene transcription. *Cell* 1989;58:1001–1007.
306. Chen J, Trounstine M, Alt FW, et al. Immunoglobulin gene rearrangement in B cell deficient mice generated by targeted deletion of the JH locus. *Int Immunol* 1993;5:647–656.
307. Grawunder U, Rolink A, Melchers F. Induction of sterile transcription from the kappa L chain gene locus in V(D)J recombinase-deficient progenitor B cells. *Int Immunol* 1995;7:1915–1925.
308. Ritchie KA, Brinster RL, Storb U. Allelic exclusion and control of endogenous immunoglobulin gene rearrangement in kappa transgenic mice. *Nature* 1984; 312:517–520.
309. Ma A, Fisher P, Dildrop R, et al. Surface IgM mediated regulation of RAG gene expression in E mu-N-myc B cell lines. *EMBO J* 1992;11:2727–2734.
310. Torres RM, Flaswinkel H, Reth M, Rajewsky K. Aberrant B cell development and immune response in mice with a compromised BCR complex. *Science* 1996;272:1804–1808.
311. Korsmeyer SJ, Hieter PA, Ravetch JV, Poplack DG, Waldmann TA, Leder P. Developmental hierarchy of immunoglobulin gene rearrangements in human leukemic pre-B-cells. *Proc Natl Acad Sci USA* 1981;78:7096–7100.
312. Zou YR, Takeda S, Rajewsky K. Gene targeting in the Ig kappa locus: efficient generation of lambda chain-expressing B cells, independent of gene rearrangements in Ig kappa. *EMBO J* 1993;12:811–820.
313. Doglio L, Kim JY, Bozek G, Storb U. Expression of lambda and kappa genes can occur in all B cells and is initiated around the same pre-B-cell developmental stage. *Dev Immunol* 1994;4:13–26.
314. Hagman J, Lo D, Doglio LT, et al. Inhibition of immunoglobulin gene rearrangement by the expression of a lambda 2 transgene. *J Exp Med* 1989;169: 1911–1929.
315. Neuberger MS, Caskey HM, Pettersson S, Williams GT, Surani MA. Isotype exclusion and transgene down-regulation in immunoglobulin-lambda transgenic mice. *Nature* 1989;338:350–352.
316. Rudin CM, Hackett JJ, Storb U. Precursors of both conventional and Ly-1 B cells can escape feedback inhibition of Ig gene rearrangement. *J Immunol* 1991;146: 3205–3210.
317. Hengstschlager M, Maizels N. Isotype exclusion in lambda 1 transgenic mice depends on transgene copy number and diminishes with down-regulation of transgene transcripts. *Eur J Immunol* 1995;25:187–191.
318. Gollahon KA, Hagman J, Brinster RL, Storb U. Ig lambda-producing B cells do not show feedback inhibition of gene rearrangement. *J Immunol* 1988;141: 2771–2780.
319. Berg J, McDowell M, Jack HM, Wabl M. Immunoglobulin lambda gene rearrangement can precede kappa gene rearrangement. *Dev Immunol* 1990;1:53–57.
320. Ramsden DA, Wu GE. Mouse kappa light-chain recombination signal sequences mediate recombination more frequently than do those of lambda light chain. *Proc Natl Acad Sci USA* 1991;88:10721–10725.
321. Arakawa H, Takeda S. Early expression of Ig mu chain from a transgene significantly reduces the duration of the pro-B stage but does not affect the small pre-B stage. *Int Immunol* 1996;8:1319–1328.
322. Harada K, Yamagishi H. Lack of feedback inhibition of V kappa gene rearrangement by productively rearranged alleles. *J Exp Med* 1991;173:409–415.
323. Levy S, Campbell MJ, Levy R. Functional immunoglobulin light chain genes are replaced by ongoing rearrangements of germline V kappa genes to downstream J kappa segment in a murine B cell line. *J Exp Med* 1989;170:1–13.
324. Kleinfield R, Hardy RR, Tarlinton D, Dangl J, Herzenberg LA, Weigert M. Recombination between an expressed immunoglobulin heavy-chain gene and a germline variable gene segment in a Ly 1+ B-cell lymphoma. *Nature* 1986;322: 843–846.
325. Radic MZ, Erikson J, Litwin S, Weigert M. B lymphocytes may escape tolerance by revising their antigen receptors. *J Exp Med* 1993;177:1165–1173.
326. Tiegs SL, Russell DM, Nemazee D. Receptor editing in self-reactive bone marrow B cells. *J Exp Med* 1993;177:1009–1020.
327. Chen C, Nagy Z, Radic MZ, et al. The site and stage of anti-DNA B-cell deletion. *Nature* 1995;373:252–255.
328. Hertz M, Nemazee D. BCR ligation induces receptor editing in IgM$^+$IgD1$^-$ bone marrow B cells in vitro. *Immunity* 1997;6:429–436.
329. Verkoczy LK, Stiernholm BJ, Berinstein NL. Up-regulation of recombination activating gene expression by signal transduction through the surface Ig receptor. *J Immunol* 1995;154:5136–5143.
330. Suzuki N, Harada T, Mihara S, Sakane T. Characterization of a germline Vk gene encoding cationic anti-DNA antibody and role of receptor editing for development of the autoantibody in patients with systemic lupus erythematosus. *J Clin Invest* 1996;98:1843–1850.
331. Han S, Zheng B, Schatz DG, Spanopoulou E, Kelsoe G. Neoteny in lymphocytes: Rag1 and Rag2 expression in germinal center B cells. *Science* 1996;274: 2094–2097.
332. Hikida M, Mori M, Takai T, Tomochika K, Hamatani K, Ohmori H. Reexpression of RAG-1 and RAG-2 genes in activated mature mouse B cells. *Science* 1996;274:2092–2094.
332a. Han S, Dillon SR, Zheng B, Shimoda M, Schlissel MS, Kelsoe G. V(D)J recombinase activity in a subset of germinal center B lymphocytes. *Science* 1997;278: 301-305.

332b. Papavasiliou F, Casellas R, Suh H, et al. V(D)J recombination in mature B cells: a mechanism for altering antibody responses. Science 1997;278: 298-301.

333. Giudicelli V, Chaume D, Bodmer J, et al. IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res* 1997;25:206-211.

334. Brodeur PH, Riblet R. The immunoglobulin heavy chain variable region (Igh-V) locus in the mouse. I. One hundred Igh-V genes comprise seven families of homologous genes. *Eur J Immunol* 1984;14:922–930.

335. Dildrop R. A new classification of mouse VH sequences. *Immunol Today* 1984; 5:85–88.

336. Meek K, Rathbun G, Reininger L, et al. Organization of the murine immunoglobulin VH complex: Placement of two new VH families (VH10 and VH11) and analysis of VH family clustering and interdigitation. *Mol Immunol* 1990;27: 1073–1081.

337. Tutter A, Brodeur P, Shlomchik M, Riblet R. Structure, map position, and evolution of two newly diverged mouse Ig VH gene families. *J Immunol* 1991;147: 3215–3223.

338. Mainville CA, Sheehan KM, Klaman LD, Giorgetti CA, Press JL, Brodeur PH. Deletional mapping of fifteen mouse VH gene families reveals a common organization for three Igh haplotypes. *J Immunol* 1996;156:1038–1046.

339. Tutter A, Riblet R. Conservation of an immunoglobulin variable-region gene family indicates a specific, noncoding function. *Proc Natl Acad Sci USA* 1989; 86:7460–7464.

340. Schroeder HWJ, Hillson JL, Perlmutter RM. Structure and evolution of mammalian VH families. *Int Immunol* 1990;2:41–50.

341. Kirkham PM, Mortari F, Newton JA, Schroeder HWJ. Immunoglobulin VH clan and family identity predicts variable domain structure and may influence antigen binding. *EMBO J* 1992;11:603–609.

342. Nei M, Gu X, Sitnikova T. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc Natl Acad Sci USA* 1997;94: 7799–7806.

343. Kofler R, Geley S, Kofler H, Helmberg A. Mouse variable-region gene families: Complexity, polymorphism and use in non-autoimmune responses. *Immunol Rev* 1992;128:5–21.

344. Livant D, Blatt C, Hood L. One heavy chain variable region gene segment subfamily in the BALB/c mouse contains 500–1000 or more members. *Cell* 1986; 47:461–470.

345. Blankenstein T, Bonhomme F, Krawinkel U. Evolution of pseudogenes in the immunoglobulin VH-gene family of the mouse. *Immunogenetics* 1987;26: 237–248.

346. Kemp DJ, Cory S, Adams JM. Cloned pairs of variable region genes for immunoglobulin heavy chains isolated from a clone library of the entire mouse genome. *Proc Natl Acad Sci USA* 1979;76:4627–4631.

347. Bothwell AL, Paskind M, Reth M, Imanishi KT, Rajewsky K, Baltimore D. Heavy chain variable region contribution to the NPb family of antibodies: Somatic mutation evident in a gamma 2a variable region. *Cell* 1981;24: 625–637.

348. Givol D, Zakut R, Effron K, Rechavi G, Ram D, Cohen JB. Diversity of germline immunoglobulin VH genes. *Nature* 1981;292:426–430.

349. Kemp DJ, Tyler B, Bernard O, et al. Organization of genes and spacers within the mouse immunoglobulin VH locus. *J Mol Appl Genet* 1981;1:245–261.

350. Rechavi G, Bienz B, Ram D, et al. Organization and evolution of immunoglobulin VH gene subgroups. *Proc Natl Acad Sci USA* 1982;79:4405–4409.

351. Rathbun GA, Capra JD, Tucker PW. Organization of the murine immunoglobulin VH complex in the inbred strains. *EMBO J* 1987;6:2931–2937.

352. Brodeur PH, Osman GE, Mackle JJ, Lalor TM. The organization of the mouse Igh-V locus. Dispersion, interspersion, and the evolution of VH gene family clusters. *J Exp Med* 1988;168:2261–2278.

353. Blankenstein T, Krawinkel U. Immunoglobulin VH region genes of the mouse are organized in overlapping clusters. *Eur J Immunol* 1987;17:1351–1357.

354. Walter MA, Dosch HM, Cox DW. A deletion map of the human immunoglobulin heavy chain variable region. *J Exp Med* 1991;174:335–349.

355. Perlmutter RM, Kearney JF, Chang SP, Hood LE. Developmentally controlled expression of immunoglobulin VH genes. *Science* 1985;227:1597–1601.

356. Malynn BA, Berman JE, Yancopoulos GD, Bona CA, Alt FW. Expression of the immunoglobulin heavy-chain variable gene repertoire. *Curr Top Microbiol Immunol* 1987;135:75–94.

357. Feeney AJ. Predominance of VH-D-JH junctions occurring at sites of short sequence homology results in limited junctional diversity in neonatal antibodies. *J Immunol* 1992;149:222–229.

358. Gauss GH, Lieber MR. The basis for the mechanistic bias for deletional over inversional V(D)J recombination. *Genes Dev* 1992;6:1553–1561.

359. Early P, Huang H, Davis M, Calame K, Hood L. An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: VH, D and JH. *Cell* 1980;19:981–992.

360. Sakano H, Kurosawa Y, Weigert M, Tonegawa S. Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes. *Nature* 1981;290:562–565.

361. Kurosawa Y, von BH, Haas W, Sakano H, Trauneker A, Tonegawa S. Identification of D segments of immunoglobulin heavy-chain genes and their rearrangement in T lymphocytes. *Nature* 1981;290:565–570.

362. Kurosawa Y, Tonegawa S. Organization, structure, and assembly of immunoglobulin heavy chain diversity DNA segments. *J Exp Med* 1982;155:201–218.

363. Feeney AJ, Riblet R. DST4: A new, and probably the last, functional DH gene in the BALB/c mouse. *Immunogenetics* 1993;37:217–221.

364. Tsubata T, Tsubata R, Reth M. Cell surface expression of the short immunoglobulin mu chain (D mu protein) in murine pre-B cells is differently regulated from that of the intact mu chain. *Eur J Immunol* 1991;21:1359–1363.

365. Haasner D, Rolink A, Melchers F. Influence of surrogate L chain on DHJH-reading frame 2 suppression in mouse precursor B cells. *Int Immunol* 1994;6:21–30.

366. Horne MC, Roth PE, DeFranco AL. Assembly of the truncated immunoglobulin heavy chain D mu into antigen receptor-like complexes in pre-B cells but not in B cells. *Immunity* 1996;4:145–158.

367. Ehlich A, Martin V, Muller W, Rajewsky K. Analysis of the B-cell progenitor compartment at the level of single cells. *Curr Biol* 1994;4:573–583.

368. Tarlinton D, Strasser A, McLean M, Basten A. DH element reading frame selection is influenced by an Ig heavy chain transgene, but not by bcl-2. *J Immunol* 1995;154:3341–3350.

369. Meek KD, Hasemann CA, Capra JD. Novel rearrangements at the immunoglobulin D locus. Inversions and fusions add to IgH somatic diversity. *J Exp Med* 1989;170:39–57.

370. Potter M, Newell JB, Rudikoff S, Haber E. Classification of mouse VK groups based on the partial amino acid sequence to the first invariant tryptophan: Impact of 14 new sequences from IgG myeloma proteins. *Mol Immunol* 1982;19: 1619–1630.

371. Kofler R, Helmberg A. A new Igk-V gene family in the mouse. *Immunogenetics* 1991;34:139–140.

372. Kofler R, Duchosal MA, Dixon FJ. Complexity, polymorphism, and connectivity of mouse Vk gene families. *Immunogenetics* 1989;29:65–74.

373. Strohal R, Helmberg A, Kroemer G, Kofler R. Mouse Vk gene classification by nucleic acid sequence similarity. *Immunogenetics* 1989;30:475–493.

374. Kroemer G, Helmberg A, Bernot A, Auffray C, Kofler R. Evolutionary relationship between human and mouse immunoglobulin kappa light chain variable region genes. *Immunogenetics* 1991;33:42–49.

375. George JB, Li S, Garrard WT. Yeast artificial chromosome contigs reveal that distal variable-region genes reside at least 3 megabases from the joining regions in the murine immunoglobulin kappa locus. *Proc Natl Acad Sci USA* 1995;92:12421–12425.

376. Kirschbaum T, Jaenichen R, Zachau HG. The mouse immunoglobulin kappa locus contains about 140 variable gene segments. *Eur J Immunol* 1996;26:1613–1620.

377. Schupp IW, Schlake T, Kirschbaum T, Zachau HG, Boehm T. A yeast artificial chromosome contig spanning the mouse immunoglobulin kappa light chain locus. *Immunogenetics* 1997;45:180–187.

378. Lee KH, Matsuda F, Kinashi T, Kodaira M, Honjo T. A novel family of variable region genes of the human immunoglobulin heavy chain. *J Mol Biol* 1987;195: 761–768.

379. Walter MA, Surti U, Hofker MH, Cox DW. The physical organization of the human immunoglobulin heavy chain gene complex. *EMBO J* 1990;9: 3303–3313.

380. Cook GP, Tomlinson IM, Walter G, et al. A map of the human immunoglobulin VH locus completed by analysis of the telomeric region of chromosome 14q. *Nature Genet* 1994;7:162–168.

381. Cook GP, Tomlinson IM. The human immunoglobulin VH repertoire. *Immunol Today* 1995;16:237–242.

382. Matsuda F, Honjo T. Organization of the human immunoglobulin heavy-chain locus. *Adv Immunol* 1996;62:1–29.

383. Schroeder HJ, Walter MA, Hofker MH, et al. Physical linkage of a human immunoglobulin heavy chain variable region gene segment to diversity and joining region elements. *Proc Natl Acad Sci USA* 1988;85:8196–8200.

384. Nagaoka H, Ozawa K, Matsuda F, et al. Recent translocation of variable and diversity segments of the human immunoglobulin heavy chain from chromosome 14 to chromosomes 15 and 16. *Genomics* 1994;22:189–197.

385. Tomlinson IM, Cook GP, Carter NP, et al. Human immunoglobulin VH and D segments on chromosomes 15q11.2 and 16p11.2. *Hum Mol Genet* 1994;3: 853–860.

386. Ravetch JV, Siebenlist U, Korsmeyer SJ, Waldmann T, Leder P. Structure of the human immunoglobulin mu locus; characterization of embryonic and rearranged J and D genes. *Cell* 1981;27:583–591.

387. Ichihara Y, Matsuoka H, Kurosawa Y. Organization of human immunoglobulin heavy chain diversity gene loci. *EMBO J* 1988;7:4141–4150.

388. Moore BB, Meek K. Recombination potential of the human DIR elements. *J Immunol* 1995;154:2175–2187.

389. Tuaillon N, Miller AB, Tucker PW, Capra JD. Analysis of direct and inverted DJH rearrangements in a human Ig heavy chain transgenic minilocus. *J Immunol* 1995;154:6453–6465.

389a. Tuaillon N, Capra JD. Use of D gene segments with irregular spacers in terminal deoxynucleotidyltransferase (TdT)+/+ and TdT-/- mice carrying a human Ig heavy chain transgenic minilocus. *Proc Natl Acad Sci U S A* 1998;95:1703-1708.

390. Nagaoka H, Ozawa K, Matsuda F, et al. Recent translocation of variable and diversity segments of the human immunoglobulin heavy chain from chromosome 14 to chromosomes 15 and 16. *Genomics* 1994;22:189–197.

391. Schable KF, Zachau HG. The variable genes of the human immunoglobulin kappa locus. *Biol Chem Hoppe Seyler* 1993;374:1001–1022.

392. Schable K, Thiebe R, Flugel A, Meindl A, Zachau HG. The human immunoglobulin kappa locus: Pseudogenes, unique and repetitive sequences. *Biol Chem Hoppe Seyler* 1994;375:189–199.

393. Ermert K, Mitlohner H, Schempp W, Zachau HG. The immunoglobulin kappa locus of primates. *Genomics* 1995;25:623–629.
394. Klein R, Zachau HG. Comparison of human germ-line kappa gene sequences to sequence data from the literature. *Eur J Immunol* 1993;23:3263–3271.
395. Arnold N, Wienberg J, Ermert K, Zachau HG. Comparative mapping of DNA probes derived from the V kappa immunoglobulin gene regions on human and great ape chromosomes by fluorescence in situ hybridization. *Genomics* 1995; 26:147–150.
396. Huber C, Thiebe R, Zachau HG. A potentially functional V kappa gene at a distance of 1.5 Mb from the immunoglobulin kappa locus. *Genomics* 1994;22: 213–215.
397. Williams SC, Frippiat JP, Tomlinson IM, Ignatovich O, Lefranc MP, Winter G. Sequence and evolution of the human germline V lambda repertoire. *J Mol Biol* 1996;264:220–232.
398. Wood DL, Coleclough C. Different joining region J elements of the murine kappa immunoglobulin light chain locus are used at markedly different frequencies. *Proc Natl Acad Sci USA* 1984;81:4756–4760.
399. Nishi M, Kataoka T, Honjo T. Preferential rearrangement of the immunoglobulin kappa chain joining region J kappa 1 and J kappa 2 segments in mouse spleen DNA. *Proc Natl Acad Sci USA* 1985;82:6399–6403.
400. Yamada M, Wasserman R, Reichard BA, Shane S, Caton AJ, Rovera G. Preferential utilization of specific immunoglobulin heavy chain diversity and joining segments in adult human peripheral blood B lymphocytes. *J Exp Med* 1991;173: 395–407.
401. Cohn M, Blomberg B, Geckler W, Raschke W, Riblet R, Weigert M. First order considerations in analyzing the generator of diversity. In: EE Sercarz, AR Williamson, CF Fox, eds. *The immune system: Genes, receptors, signals.* New York: Academic Press, 1974;89.
402. Weigert M, Riblet R. Genetic control of antibody variable regions. *Cold Spring Harb Symp Quant Biol* 1977;2:837–846.
403. Gearhart PJ, Johnson ND, Douglas R, Hood L. IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts. *Nature* 1981;291: 29–34.
404. Crews S, Griffin J, Huang H, Calame K, Hood L. A single VH gene segment encodes the immune response to phosphorylcholine: Somatic mutation is correlated with the class of the antibody. *Cell* 1981;25:59–66.
405. Manser T, Parhami SB, Margolies MN, Gefter ML. Somatically mutated forms of a major anti-p-azophenylarsonate antibody variable region with drastically reduced affinity for p-azophenylarsonate. By-products of an antigen-driven immune response? *J Exp Med* 1987;166:1456–1463.
406. Ray SK, Putterman C, Diamond B. Pathogenic autoantibodies are routinely generated during the response to foreign antigen: A paradigm for autoimmune disease. *Proc Natl Acad Sci USA* 1996;93:2019–2024.
407. Liu YJ, Joshua DE, Williams GT, Smith CA, Gordon J, MacLennan IC. Mechanism of antigen-driven selection in germinal centres. *Nature* 1989;342:929–931.
408. Shokat KM, Goodnow CC. Antigen-induced B-cell death and elimination during germinal-centre immune responses. *Nature* 1995;375:334–338.
409. Choe J, Kim HS, Zhang X, Armitage RJ, Choi YS. Cellular and molecular factors that regulate the differentiation and apoptosis of germinal center B cells. Anti-Ig down-regulates Fas expression of CD40 ligand-stimulated germinal center B cells and inhibits Fas-mediated apoptosis. *J Immunol* 1996;157:1006–1016.
410. Eisen HN, Siskind GW. Variations in affinities of antibodies during the immune response. *Biochemistry* 1964;3:996–1008.
411. Jacob J, Kelsoe G, Rajewsky K, Weiss U. Intraclonal generation of antibody mutants in germinal centres. *Nature* 1991;354:389–392.
412. Jacob J, Kelsoe G. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for periarteriolar lymphoid sheath-associated foci and germinal centers. *J Exp Med* 1992;176:679–687.
413. Pascual V, Liu YJ, Magalski A, de Boutellier O, Banchereau J, Capra JD. Analysis of somatic mutation in five B cell subsets of human tonsil. *J Exp Med* 1994; 180:329–339.
414. Lebecque S, de Boutellier O, Arpin C, Banchereau J, Liu YJ. Germinal center founder cells display propensity for apoptosis before onset of somatic mutation. *J Exp Med* 1997;185:563–571.
415. Zheng B, Han S, Kelsoe G. T helper cells in murine germinal centers are antigen-specific emigrants that downregulate Thy-1. *J Exp Med* 1996;184:1083–1091.
416. Razanajaona D, van Kooten C, Lebecque S, et al. Somatic mutations in human Ig variable genes correlate with a partially functional CD40-ligand in the X-linked hyper-IgM syndrome. *J Immunol* 1996;157:1492–1498.
417. Casamayor PM, Khan M, MacLennan IC. A subset of CD4$^+$ memory T cells contains preformed CD40 ligand that is rapidly but transiently expressed on their surface after activation through the T cell receptor complex. *J Exp Med* 1995;181:1293–1301.
418. Han S, Zheng B, Dal PJ, Kelsoe G. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. IV. Affinity-dependent, antigen-driven B cell apoptosis in germinal centers as a mechanism for maintaining self-tolerance. *J Exp Med* 1995;182:1635–1644.
419. Pulendran B, Kannourakis G, Nouri S, Smith KG, Nossal GJ. Soluble antigen can cause enhanced apoptosis of germinal-centre B cells. *Nature* 1995;375: 331–334.
419a. Hande S, Notidis E, Manser T. Bcl-2 obstructs negative selection of autoreactive, hypermutated antibody V regions during memory B cell development. *Immunity* 1998;8:189-198.
420. Kuppers R, Zhao M, Hansmann ML, Rajewsky K. Tracing B cell development in human germinal centres by molecular analysis of single cells picked from histological sections. *EMBO J* 1993;12:4955–4967.
421. Decker DJ, Linton PJ, Zaharevitz S, Biery M, Gingeras TR, Klinman NR. Defining subsets of naive and memory B cells based on the ability of their progeny to somatically mutate in vitro. *Immunity* 1995;2:195–203.
422. Kepler TB, Perelson AS. Cyclic re-entry of germinal center B cells and the efficiency of affinity maturation. *Immunol Today* 1993;14:412–415.
422a. Matsumoto M, Lo SF, Carruthers CJ, et al. Affinity maturation without germinal centres in lymphotoxin-alpha-deficient mice. *Nature* 1996;382:462-466.
423. Kim S, David M, Sinn E, Patten P, Hood L. Antibody diversity: Somatic hypermutation of rearranged VH genes. *Cell* 1981;27:573–581.
424. Lebecque SG, Gearhart PJ. Boundaries of somatic mutation in rearranged immunoglobulin genes: 5′ boundary is near the promoter, and 3′ boundary is approximately 1 kb from V(D)J gene. *J Exp Med* 1990;172:1717–1727.
425. Gorski J, Rollini P, Mach B. Somatic mutations of immunoglobulin variable genes are restricted to the rearranged V gene. *Science* 1983;220:1179–1181.
426. Sablitzky F, Weisbaum D, Rajewsky K. Sequence analysis of non-expressed immunoglobulin heavy chain loci in clonally related, somatically mutated hybridoma cells. *EMBO J* 1985;4:3435–3437.
427. Giusti AM, Manser T. Hypermutation is observed only in antibody H chain V region transgenes that have recombined with endogenous immunoglobulin H DNA: Implications for the location of cis-acting elements required for somatic mutation. *J Exp Med* 1993;177:797–809.
428. Weiss S, Wu GE. Somatic point mutations in unrearranged immunoglobulin gene segments encoding the variable region of lambda light chains. *EMBO J* 1987;6:927–932.
429. Picard D, Schaffner W. Unrearranged immunoglobulin lambda variable region is transcribed in kappa-producing myelomas. *EMBO J* 1984;3:3031–3035.
430. O'Brien RL, Brinster RL, Storb U. Somatic hypermutation of an immunoglobulin transgene in kappa transgenic mice. *Nature* 1987;326:405–409.
431. Sharpe MJ, Milstein C, Jarvis JM, Neuberger MS. Somatic hypermutation of immunoglobulin kappa may depend on sequences 3′ of C kappa and occurs on passenger transgenes. *EMBO J* 1991;10:2139–2145.
432. Betz AG, Milstein C, Gonzalez FA, Pannell R, Larson T, Neuberger MS. Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region. *Cell* 1994;77: 239–248.
432a. Goyenechea B, Klix N, Yelamos J, et al. Cells strongly expressing Ig(kappa) transgenes show clonal recruitment of hypermutation: a role for both MAR and the enhancers. *EMBO J* 1997;16:3987-3994.
432b. Klix N, Jolly CJ, Davies SL, Bruggemann M, Williams GT, Neuberger MS. Multiple sequences from downstream of the J kappa cluster can combine to recruit somatic hypermutation to a heterologous, upstream mutation domain. *Eur J Immunol* 1998;28:317-326.
433. Peters A, Storb U. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity* 1996;4:57–65.
434. Bachl J, Wabl M. Enhancers of hypermutation. *Immunogenetics* 1996;45:59–64.
435. Yelamos J, Klix N, Goyenechea B, et al. Targeting of non-Ig sequences in place of the V segment by somatic hypermutation. *Nature* 1995;376:225–229.
436. Tumas-Brundage K, Vora KA, Giusti AM, Manser T. Characterization of the cis-acting elements required for somatic hypermutation of murine antibody V genes using conventional transgenic and transgene homologous recombination approaches. *Semin Immunol* 1996;8:141–150.
437. Hengstschlager M, Williams M, Maizels N. A lambda 1 transgene under the control of a heavy chain promoter and enhancer does not undergo somatic hypermutation. *Eur J Immunol* 1994;24:1649–1656.
438. Storb U. The molecular basis of somatic hypermutation of immunoglobulin genes. *Curr Opin Immunol* 1996;8:206–214.
438a. Migliazza A, Martinotti S, Chen W, et al. Frequent somatic hypermutation of the 5′ noncoding region of the BCL6 gene in B-cell lymphoma. *Proc Natl Acad Sci USA* 1995;92:12520–12524.
439. Goyenechea B, Milstein C. Modifying the sequence of an immunoglobulin V-gene alters the resulting pattern of hypermutation. *Proc Natl Acad Sci USA* 1996; 93:13979–13984.
440. Dorner T, Brezinschek HP, Brezinschek RI, Foster SJ, Domiati SR, Lipsky PE. Analysis of the frequency and pattern of somatic mutations within nonproductively rearranged human variable heavy chain genes. *J Immunol* 1997;158: 2779–2789.
441. Bachl J, Wabl M. An immunoglobulin mutator that targets G.C base pairs. *Proc Natl Acad Sci USA* 1996;93:851–855.
442. Betz AG, Rada C, Pannell R, Milstein C, Neuberger MS. Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: Clustering, polarity, and specific hot spots. *Proc Natl Acad Sci USA* 1993;90: 2385–2388.
443. Insel RA, Varade WS. Bias in somatic hypermutation of human VH genes. *Int Immunol* 1994;6:1437–1443.
444. Dildrop R, Bruggemann M, Radbruch A, Rajewsky K, Beyreuther K. Immunoglobulin V region variants in hybridoma cells. II. Recombination between V genes. *EMBO J* 1982;1:635–640.
445. Xu B, Selsing E. Analysis of sequence transfers resembling gene conversion in a mouse antibody transgene. *Science* 1994;265:1590–1593.
446. Reynaud CA, Anquez V, Dahan A, Weill JC. A single rearrangement event gen-

erates most of the chicken immunoglobulin light chain diversity. *Cell* 1985;40: 283–291.
447. Becker RS, Knight KL. Somatic diversification of immunoglobulin heavy chain VDJ genes: evidence for somatic gene conversion in rabbits. *Cell* 1990;63: 987–997.
448. Sun J, Butler JE. Molecular characterization of VDJ transcripts from a newborn piglet. *Immunology* 1996;88:331–339.
449. Bentley DL, Rabbitts TH. Evolution of immunoglobulin V genes: evidence indicating that recently duplicated human V kappa sequences have diverged by gene conversion. *Cell* 1983;32:181–189.
450. Cohen JB, Givol D. Allelic immunoglobulin VH genes in two mouse strains: Possible germline gene recombination. *EMBO J* 1983;2:2013–2018.
451. Rogerson BJ. Somatic hypermutation of VHS107 genes is not associated with gene conversion among family members. *Int Immunol* 1995;7:1225–1235.
452. Golding GB, Gearhart PJ, Glickman BW. Patterns of somatic mutations in immunoglobulin variable genes. *Genetics* 1987;115:169–176.
453. Seidman MM, Bredberg A, Seetharam S, Kraemer KH. Multiple point mutations in a shuttle vector propagated in human cells: Evidence for an error-prone DNA polymerase activity. *Proc Natl Acad Sci USA* 1987;84:4944–4948.
454. Kim N, Kage K, Matsuda F, Lefranc MP, Storb U. B lymphocytes of xeroderma pigmentosum or Cockayne syndrome patients with inherited defects in nucleotide excision repair are fully capable of somatic hypermutation of immunoglobulin genes. *J Exp Med* 1997;186:413–419.
454a. Shen HM, Cheo DL, Friedberg E, Storb U. The inactivation of the XP-C gene does not affect somatic hypermutation or class switch recombination of immunoglobulin genes. *Mol Immunol* 1997;34:527-533.
454b. Rada C, Yelamos J, Dean W, Milstein C. The 5' hypermutation boundary of kappa chains is independent of local and neighboring sequences and related to the distance from the initiation of transcription. *Eur J Immunol* 1997;27:3115-3120.
454c. Tumas-Brundage K, Manser T. The transcriptional promoter regulates hypermutation of the antibody heavy chain locus. *J Exp Med* 1997;185:239-250.
454d. Cascalho M, Wong J, Steinberg C, Wabl M. Mismatch repair co-opted by hypermutation. *Science* 1998;279:1207-1210.
454e. Narayanan L, Fritzell MA, Baker SM, Liskay RM, Glazer PM. Elevated levels of mutation in multiple tissues of mice deficient in the DNA mismatch repair gene Pms2. *Proc Natl Acad Sci USA* 1997;94:3122-3127.
454f. Razanajaona D, Denepoux S, Blanchard D, et al. In vitro triggering of somatic mutation in human naive B cells. *J Immunol* 1997;159:3347-3353.
455. Roes J, Rajewsky K. Immunoglobulin D (IgD)-deficient mice reveal an auxiliary receptor function for IgD in antigen-mediated recruitment of B cells. *J Exp Med* 1993;177:45–55.
456. Liu YJ, de Boutellier O, Arpin C, et al. Normal human IgD$^+$IgM$^-$ germinal center B cells can express up to 80 mutations in the variable region of their IgD transcripts. *Immunity* 1996;4:603–613.
457. Kalliberg E, Jainandunsing S, Gray D, Leanderson T. Somatic mutation of immunoglobulin V genes in vitro. *Science* 1996;271:1285–1289.
458. Denepoux S, Razanajaona D, Blanchard D, et al. Induction of somatic mutation in a human B cell line in vitro. *Immunity* 1997;6:35–46.
459. Jolly CJ, Klix N, Neuberger MS. Rapid methods for the analysis of immunoglobulin gene hypermutation: Application to transgenic and gene targeted mice. *Nucleic Acids Res* 1997;25:1913–1919.
460. Williams AF. A year in the life of the immunoglobulin superfamily. *Immunol Today* 1987;8:298–303.
461. Hunkapiller T, Hood L. Diversity of the immunoglobulin gene superfamily. *Adv Immunol* 1989;44:1–63.
462. Matsunaga T, Mori N. The origin of the immune system. The possibility that immunoglobulin superfamily molecules and cell adhesion molecules of chicken and slime mould are all related. *Scand J Immunol* 1987;25:485–495.
463. Hoek RM, Smit AB, Frings H, et al. A new Ig-superfamily member, molluscan defence molecule (MDM) from *Lymnaea stagnalis*, is down-regulated during parasitosis. *Eur J Immunol* 1996;26:939–944.
464. Sun SC, Lindstrom I, Boman HG, Faye I, Schmidt O. Hemolin: An insect-immune protein belonging to the immunoglobulin superfamily. *Science* 1990; 250:1729–1732.
465. Lindstrom-Dinnetz I, Sun SC, Faye I. Structure and expression of Hemolin, an insect member of the immunoglobulin gene superfamily. *Eur J Biochem* 1995; 230:920–925.
466. Parnes JR, Hunkapiller T. L3T4 and the immunoglobulin gene superfamily: New relationships between the immune system and the nervous system. *Immunol Rev* 1987;100:109–127.
467. Chretien I, Robert J, Marcuz A, Garcia SJ, Courtet M, Du PL. CTX, a novel molecule specifically expressed on the surface of cortical thymocytes in Xenopus. *Eur J Immunol* 1996;26:780–791.
468. Johnson P, Williams AF. Striking similarities between antigen receptor J pieces and sequence in the second chain of the murine CD8 antigen. *Nature* 1986;323:74–76.
469. Siu G, Kronenberg M, Strauss E, Haars R, Mak TW, Hood L. The structure, rearrangement and expression of D beta gene segments of the murine T-cell antigen receptor. *Nature* 1984;311:344–350.
470. Rast JP, Litman GW. T-cell receptor gene homologs are present in the most primitive jawed vertebrates. *Proc Natl Acad Sci USA* 1994;91:9248–9252.
471. Hawke NA, Rast JP, Litman GW. Extensive diversity of transcribed TCR-beta in phylogenetically primitive vertebrate. *J Immunol* 1996;156:2458–2464.
472. Rast JP, Anderson MK, Strong SJ, Luer C, Litman RT, Litman GW. Alpha, beta, gamma, and delta T cell antigen receptor genes arose early in vertebrate phylogeny. *Immunity* 1997;6:1–11.
473. Hinds KR, Litman GW. Major reorganization of immunoglobulin VH segmental elements during vertebrate evolution. *Nature* 1986;320:546–549.
474. Reynaud CA, Anquez V, Grimal H, Weill JC. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell* 1987;48:379–388.
475. Thompson CB, Neiman PE. Somatic diversification of the chicken immunoglobulin light chain gene is limited to the rearranged variable gene segment. *Cell* 1987;48:369–378.
476. Knight KL, Becker RS. Molecular basis of the allelic inheritance of rabbit immunoglobulin VH allotypes: Implications for the generation of antibody diversity. *Cell* 1990;60:963–970.
476a. Lanning DK, Knight KL. Somatic hypermutation: mutations 3' of rabbit VDJ H-chain genes. *J Immunol* 1997;159:4403-4407.
477. Parng CL, Hansal S, Goldsby RA, Osborne BA. Gene conversion contributes to Ig light chain diversity in cattle. *J Immunol* 1996;157:5478–5486.
478. Jack HM, Wabl M. Immunoglobulin mRNA stability varies during B lymphocyte differentiation. *EMBO J* 1988;7:1041–1046.
479. Bode J, Kohwi Y, Dickinson L, et al. Biological significance of unwinding capability of nuclear matrix-associating DNAs. *Science* 1992;255:195–197.
480. Freeman LA, Garrard WT. DNA supercoiling in chromatin structure and gene expression. *Crit Rev Eukaryot Gene Expr* 1992;2:165–209.
481. Baron MH. Developmental regulation of the vertebrate globin multigene family. *Gene Expr* 1996;6:129–137.
482. Razin A, Riggs AD. DNA metylation and gene function. *Science* 1980;210: 604–610.
483. Rogers J, Wall R. Immunoglobulin heavy chain genes: Demethylation accompanies class switching. *Proc Natl Acad Sci USA* 1981;78:7497–7501.
484. Storb U, Wilson R, Selsing E, Walfield A. Rearranged and germline immunoglobulin kappa genes: Different states of DNase I sensitivity of constant kappa genes in immunocompetent and nonimmune cells. *Biochemistry* 1981;20:990–996.
485. Mather EL, Perry RP. Methylation status and DNase I sensitivity of immunoglobulin genes: Changes associated with rearrangement. *Proc Natl Acad Sci USA* 1983;80:4689–4693.
486. Mills FC, Fisher LM, Kuroda R, Ford AM, Gould HJ. DNase I hypersensitive sites in the chromatin of human mu immunoglobulin heavy-chain genes. *Nature* 1983;306:809–812.
487. Blackman MA, Koshland ME. Specific 5' and 3' regions of the mu chain gene are undermethylated at distinct stages of B-cell differentiation. *Proc Natl Acad Sci USA* 1985;82:3809–3813.
488. Stavnezer-Nordgren J, Sirlin S. Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. *EMBO J* 1986;5:95–102.
489. Pfeiffer W, Zachau HG. Accessibility of expressed and non-expressed genes to a restriction nuclease. *Nucleic Acids Res* 1980;8:4621–4638.
490. Madisen L, Groudine M. Identification of a locus control region in the immunoglobulin heavy-chain locus that deregulates c-myc expression in plasmacytoma and Burkitt's lymphoma cells. *Genes Dev* 1994;8:2212–2226.
491. Mills FC, Harindranath N, Mitchell M, Max EE. Enhancer complexes located downstream of both human immunoglobulin C-alpha genes. *J Exp Med* 1997; 186:845–858.
492. Weischet WO, Glotov BO, Schnell H, Zachau HG. Differences in the nuclease sensitivity between the two alleles of the immunoglobulin kappa light chain genes in mouse liver and myeloma nuclei. *Nucleic Acids Res* 1982;10: 3627–3645.
493. Gimble JM, Max EE. Human immunoglobulin kappa gene enhancer: Chromatin structure analysis at high resolution. *Mol Cell Biol* 1987;7:15–25.
494. Singh H, LeBowitz JH, Baldwin AJ, Sharp PA. Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA. *Cell* 1988;52:415–423.
495. Falkner FG, Zachau HG. Correct transcription of an immunoglobulin kappa gene requires an upstream fragment containing conserved sequence elements. *Nature* 1984;310:71–74.
496. Parslow TG, Blair DL, Murphy WJ, Granner DK. Structure of the 5' ends of immunoglobulin genes: A novel conserved sequence. *Proc Natl Acad Sci USA* 1984;91:2650–2654.
497. Bergman Y, Rice D, Grosschedl R, Baltimore D. Two regulatory elements for immunoglobulin kappa light chain gene expression. *Proc Natl Acad Sci USA* 1984;81:7041–7045.
498. Grosschedl R, Baltimore D. Cell-type specificity of immunoglobulin gene expression is regulated by at least three DNA sequence elements. *Cell* 1985;41: 885–897.
499. Picard D, Schaffner W. Cell-type preference of immunoglobulin kappa and lambda gene promoters. *EMBO J* 1985;4:2831–2838.
500. Ballard DW, Bothwell A. Mutational analysis of the immunoglobulin heavy chain promoter region. *Proc Natl Acad Sci USA* 1986;83:9626–9630.
501. Wirth T, Staudt L, Baltimore D. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. *Nature* 1987;329: 174–178.
502. Mizushima-Sugano J, Roeder RG. Cell-type-specific transcription of an immunoglobulin kappa light chain gene in vitro. *Proc Natl Acad Sci USA* 1986;83: 8511–8515.
503. Hermanson GG, Briskin M, Sigman D, Wall R. Immunoglobulin enhancer and

promoter motifs 5' of the B29 B-cell- specific gene. *Proc Natl Acad Sci USA* 1989;86:7341–7345.
504. Christensen SM, Martin BK, Tan SS, Weis JH. Identification of sites for distinct DNA binding proteins including Oct-1 and Oct-2 in the Cr2 gene. *J Immunol* 1992;148:3610–3617.
505. Thevenin C, Lucas BP, Kozlow EJ, Kehrl JH. Cell type- and stage-specific expression of the CD20/B1 antigen correlates with the activity of a diverged octamer DNA motif present in its promoter. *J Biol Chem* 1993;268:5949–5956.
506. Annweiler A, Muller IM, Wirth T. Oct2 transactivation from a remote enhancer position requires a B-cell-restricted activity. *Mol Cell Biol* 1992;12:3107–3116.
507. Clerc RG, Corcoran LM, LeBowitz JH, Baltimore D, Sharp PA. The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. *Genes Dev* 1988;2:1570–1581.
508. Ko HS, Fast P, McBride W, Staudt LM. A human protein specific for the immunoglobulin octamer DNA motif contains a functional homeobox domain. *Cell* 1988;55:135–144.
509. Staudt LM, Clerc RG, Singh H, LeBowitz JH, Sharp PA, Baltimore D. Cloning of a lymphoid-specific cDNA encoding a protein binding the regulatory octamer DNA motif. *Science* 1988;241:577–580.
510. Sturm RA, Das G, Herr W. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes Dev* 1988;2: 1582–1599.
511. Klemm JD, Rould MA, Aurora R, Herr W, Pabo CO. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* 1994;77:21–32.
512. Gerster T, Balmaceda CG, Roeder RG. The cell type-specific octamer transcription factor OTF-2 has two domains required for the activation of transcription. *EMBO J* 1990;9:1635–1643.
513. Tanaka M, Herr W. Differential transcriptional activation by Oct-1 and Oct-2: Interdependent activation domains induce Oct-2 phosphorylation. *Cell* 1990;60: 375–386.
514. Yang J, Muller IM, Seipel K, et al. Both Oct-1 and Oct-2A contain domains which can activate the ubiquitously expressed U2 snRNA genes. *EMBO J* 1991; 10:2291–2296.
515. Tanaka M, Lai JS, Herr W. Promoter-selective activation domains in Oct-1 and Oct-2 direct differential activation of an snRNA and mRNA promoter. *Cell* 1992;68:755–767.
516. Feldhaus AL, Klug CA, Arvin KL, Singh H. Targeted disruption of the Oct-2 locus in a B cell provides genetic evidence for two distinct cell type-specific pathways of octamer element-mediated gene activation. *EMBO J* 1993;12:2763–2772.
517. Corcoran LM, Karvelas M, Nossal GJ, Ye ZS, Jacks T, Baltimore D. Oct-2, although not required for early B-cell development, is critical for later B-cell maturation and for postnatal survival. *Genes Dev* 1993;7:570–582.
518. Corcoran LM, Karvelas M. Oct-2 is required early in T cell-independent B cell activation for G1 progression and for proliferation. *Immunity* 1994;1:635–645.
519. Pierani A, Heguy A, Fujii H, Roeder RG. Activation of octamer-containing promoters by either octamer-binding transcription factor 1 (OTF-1) or OTF-2 and requirement of an additional B-cell-specific component for optimal transcription of immunoglobulin promoters. *Mol Cell Biol* 1990;10:6204–6215.
520. Luo Y, Roeder RG. Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator OCA-B. *Mol Cell Biol* 1995;15:4115–4124.
521. Strubin M, Newell JW, Matthias P. OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. *Cell* 1995;80:497–506.
522. Gstaiger M, Knoepfel L, Georgiev O, Schaffner W, Hovens CM. A B-cell coactivator of octamer-binding transcription factors. *Nature* 1995;373:360–362.
523. Gstaiger M, Georgiev O, van LH, et al. The B cell coactivator Bob1 shows DNA sequence-dependent complex formation with Oct-1/Oct-2 factors, leading to differential promoter activation. *EMBO J* 1996;15:2781–2790.
524. Cepek KL, Chasman DI, Sharp PA. Sequence-specific DNA binding of the B-cell-specific coactivator OCA-B. *Genes Dev* 1996;10:2079–2088.
525. Kim U, Qin XF, Gong S, et al. The B-cell-specific transcription coactivator OCA-B/OBF-1/Bob-1 is essential for normal production of immunoglobulin isotypes. *Nature* 1996;383:542–547.
526. Schubart DB, Rolink A, Kosco VM, Botteri F, Matthias P. B-cell-specific coactivator OBF-1/OCA-B/Bob1 required for immune response and germinal centre formation. *Nature* 1996;383:538–542.
527. Luo Y, Fujii H, Gerster T, Roeder RG. A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors. *Cell* 1992;71:231–241.
528. Pfisterer P, Annweiler A, Ullmer C, Corcoran LM, Wirth T. Differential transactivation potential of Oct1 and Oct2 is determined by additional B cell-specific activities. *EMBO J* 1994;13:1655–1663.
529. Cook GP, Neuberger MS. Lymphoid-specific transcriptional activation by components of the IgH enhancer: Studies on the E2/E3 and octanucleotide elements. *Nucleic Acids Res* 1990;18:3565–3571.
530. Stern S, Tanaka M, Herr W. The Oct-1 homoeodomain directs formation of a multiprotein-DNA complex with the HSV transactivator VP16. *Nature* 1989; 341:624–630.
531. Kristie TM, Sharp PA. Interactions of the Oct-1 POU subdomains with specific DNA sequences and with the HSV alpha-trans-activator protein. *Genes Dev* 1990;4:2383–2396.
532. Murphy S, Yoon JB, Gerster T, Roeder RG. Oct-1 and Oct-2 potentiate functional interactions of a transcription factor with the proximal sequence element of small nuclear RNA genes. *Mol Cell Biol* 1992;12:3247–3261.
532a. Pelletier MR, Hatada EN, Scholz G, Scheidereit C. Efficient transcription of an immunoglobulin kappa promoter requires specific sequence elements overlapping with and downstream of the transcriptional start site. *Nucl Acids Res* 1997; 25:3995-4003.
533. Lenardo M, Pierce JW, Baltimore D. Protein-binding sites in Ig gene enhancers determine transcriptional activity and inducibility. *Science* 1987;236:1573–1577.
534. Jenuwein T, Grosschedl R. Complex pattern of immunoglobulin mu gene expression in normal and transgenic mice: Nonoverlapping regulatory sequences govern distinct tissue specificities. *Genes Dev* 1991;5:932–943.
535. Yuan D, Dang T, Hawley J, Jenuwein T, Grosschedl R. Role of the OCTA site in regulation of IgH chain gene transcription during B cell activation. *Int Immunol* 1995;7:1163–1172.
536. Su LK, Kadesch T. The immunoglobulin heavy-chain enhancer functions as the promoter for I mu sterile transcription. *Mol Cell Biol* 1990;10:2619–2624.
537. Currie RA, Roeder RG. Identification of an octamer-binding site in the mouse kappa light-chain immunoglobulin enhancer. *Mol Cell Biol* 1989;9:4239–4247.
538. Kemler I, Schaffner W. Octamer transcription factors and the cell type-specificity of immunoglobulin gene expression. *FASEB J* 1990;4:1444–1449.
539. Wirth T, Priess A, Annweiler A, Zwilling S, Oeler B. Multiple Oct2 isoforms are generated by alternative splicing. *Nucleic Acids Res* 1991;19:43–51.
540. Stoykova AS, Sterrer S, Erselius JR, Hatzopoulos AK, Gruss P. Mini-Oct and Oct-2c: Two novel, functionally diverse murine Oct-2 gene products are differentially expressed in the CNS. *Neuron* 1992;8:541–558.
541. Segil N, Roberts SB, Heintz N. Mitotic phosphorylation of the Oct-1 homeodomain and regulation of Oct-1 DNA binding activity. *Science* 1991;254: 1814–1816.
542. Zwilling S, Konig H, Wirth T. High mobility group protein 2 functionally interacts with the POU domains of octamer transcription factors. *EMBO J* 1995;14: 1198–1208.
543. Fontes JD, Jabrane FN, Toth CR, Peterlin BM. Binding and cooperative interactions between two B cell-specific transcriptional coactivators. *J Exp Med* 1996; 183:2517–2521.
544. Atchison ML, Delmas V, Perry RP. A novel upstream element compensates for an ineffectual octamer motif in an immunoglobulin V kappa promoter. *EMBO J* 1990;9:3109–3117.
545. Sigvardsson M, Akerblad P, Leanderson T. Early B cell factor interacts with a subset of kappa promoters. *J Immunol* 1996;156:3788–3796.
546. Eaton S, Calame K. Multiple DNA sequence elements are necessary for the function of an immunoglobulin heavy chain promoter. *Proc Natl Acad Sci USA* 1987;84:7634–7638.
547. Landolfi NF, Yin XM, Capra JD, Tucker PW. A conserved heptamer upstream of the IgH promoter region octamer can be the site of a coordinate protein-DNA interaction. *Nucleic Acids Res* 1988;16:5503–5514.
548. Kemler I, Schreiber E, Muller MM, Matthias P, Schaffner W. Octamer transcription factors bind to two different sequence motifs of the immunoglobulin heavy chain promoter. *EMBO J* 1989;8:2001–2008.
549. LeBowitz JH, Clerc RG, Brenowitz M, Sharp PA. The Oct-2 protein binds cooperatively to adjacent octamer sites. *Genes Dev* 1989;3:1625–1638.
550. Poellinger L, Roeder RG. Octamer transcription factors 1 and 2 each bind to two different functional elements in the immunoglobulin heavy-chain promoter. *Mol Cell Biol* 1989;9:747–756.
551. Poellinger L, Yoza BK, Roeder RG. Functional cooperativity between protein molecules bound at two distinct sequence elements of the immunoglobulin heavy-chain promoter. *Nature* 1989;337:573–576.
552. Yoza BK, Roeder RG. Identification of a novel factor that interacts with an immunoglobulin heavy-chain promoter and stimulates transcription in conjunction with the lymphoid cell-specific factor OTF2. *Mol Cell Biol* 1990;10: 2145–2153.
553. Cooper C, Johnson D, Roman C, Avitahl N, Tucker P, Calame K. The C/EBP family of transcriptional activators is functionally important for Ig VH promoter activity in vivo and in vitro. *J Immunol* 1992;149:3225–3231.
554. Webb CF, Das C, Eaton S, Calame K, Tucker PW. Novel protein-DNA interactions associated with increased immunoglobulin transcription in response to antigen plus interleukin-5. *Mol Cell Biol* 1991;11:5197–5205.
555. Herrscher RF, Kaplan MH, Lelsz DL, Das C, Scheuermann R, Tucker PW. The immunoglobulin heavy-chain matrix-associating regions are bound by Bright: A B cell-specific trans-activator that describes a new DNA-binding protein family. *Genes Dev* 1995;9:3067–3082.
556. Avitahl N, Calame K. A 125 bp region of the Ig VH1 promoter is sufficient to confer lymphocyte-specific expression in transgenic mice. *Int Immunol* 1996;8: 1359–1366.
557. Buchanan KL, Hodgetts SI, Byrnes J, Webb CF. Differential transcription efficiency of two Ig VH promoters in vitro. *J Immunol* 1995;155:4270–4277.
558. Stiernholm NB, Berinstein NL. A mutated promoter of a human Ig V lambda gene segment is associated with reduced germ-line transcription and a low frequency of rearrangement. *J Immunol* 1995;154:1748–1761.
559. Buchanan KL, Smith EA, Dou S, Corcoran LM, Webb CF. Family-specific differences in transcription efficiency of Ig heavy chain promoters. *J Immunol* 1997;159:1247–1254.
560. Sawadogo M, Roeder RG. Interaction of a gene-specific transcription factor with

the adenovirus major late promoter upstream of the TATA box region. *Cell* 1985;43:165–175.
561. Chang LA, Smith T, Pognonec P, Roeder RG, Murialdo H. Identification of USF as the ubiquitous murine factor that binds to and stimulates transcription from the immunoglobulin lambda 2-chain promoter. *Nucleic Acids Res* 1992;20:287–293.
562. Blackwell TK, Moore MW, Yancopoulos GD, et al. Recombination between immunoglobulin variable region gene segments is enhanced by transcription. *Nature* 1986;324:585–589.
563. Berman JE, Humphries CG, Barth J, Alt FW, Tucker PW. Structure and expression of human germline VH transcripts. *J Exp Med* 1991;173:1529–1535.
564. Perry RP, Kelley DE, Coleclough C, et al. Transcription of mouse kappa chain genes: Implications for allelic exclusion. *Proc Natl Acad Sci USA* 1980;77: 1937–1941.
565. Martin D, Huang RQ, LeBien T, Van Ness B. Induced rearrangement of kappa genes in the BLIN-1 human pre-B cell line correlates with germline J-C kappa and V kappa transcription. *J Exp Med* 1991;173:639–645.
566. Nelson KJ, Haimovich J, Perry RP. Characterization of productive and sterile transcripts from the immunoglobulin heavy-chain locus: Processing of micron and muS mRNA. *Mol Cell Biol* 1983;3:1317–1332.
567. Lennon GG, Perry RP. C mu-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5′-nontranslatable exon. *Nature* 1985;318:475–478.
568. Neale GA, Kitchingman GR. mRNA transcripts initiating within the human immunoglobulin mu heavy chain enhancer region contain a non-translatable exon and are extremely heterogeneous at the 5′ end. *Nucleic Acids Res* 1991;19: 2427–2433.
569. Li SC, Rothman PB, Zhang J, Chan C, Hirsh D, Alt FW. Expression of I mu-C gamma hybrid germline transcripts subsequent to immunoglobulin heavy chain class switching. *Int Immunol* 1994;6:491–497.
570. Alessandrini A, Desiderio SV. Coordination of immunoglobulin DJH transcription and D-to-JH rearrangement by promoter-enhancer approximation. *Mol Cell Biol* 1991;11:2096–2107.
571. Kottmann AH, Brack C, Eibel H, Kohler G. A survey of protein-DNA interaction sites within the murine immunoglobulin heavy chain locus reveals a particularly complex pattern around the DQ52 element. *Eur J Immunol* 1992;22:2113–2120.
572. Kottmann AH, Zevnik B, Welte M, Nielsen PJ, Kohler G. A second promoter and enhancer element within the immunoglobulin heavy chain locus. *Eur J Immunol* 1994;24:817–821.
573. Martin DJ, Van Ness B. Initiation and processing of two kappa immunoglobulin germ line transcripts in mouse B cells. *Mol Cell Biol* 1990;10:1950–1958.
574. Weaver D, Baltimore D. B lymphocyte-specific protein binding near an immunoglobulin kappa-chain gene J segment. *Proc Natl Acad Sci USA* 1987;84: 1516–1520.
575. Frances V, Pandrau GD, Guret C, et al. A surrogate 15 kDa JC kappa protein is expressed in combination with mu heavy chain by human B cell precursors. *EMBO J* 1994;13:5937–5943.
576. Daitch LE, Moore MW, Persiani DM, Durdik JM, Selsing E. Transcription and recombination of the murine RS element. *J Immunol* 1992;149:832–840.
577. Kuhn R, Rajewsky K, Muller W. Generation and analysis of interleukin-4 deficient mice. *Science* 1991;254:707–710.
578. Rothman P, Chen YY, Lutzker S, et al. Structure and expression of germ line immunoglobulin heavy-chain epsilon transcripts: Interleukin-4 plus lipopolysaccharide-directed switching to C epsilon. *Mol Cell Biol* 1990;10:1672–1679.
579. Delphin S, Stavnezer J. Characterization of an interleukin 4 (IL-4) responsive region in the immunoglobulin heavy chain germline epsilon promoter: Regulation by NF-IL-4, a C/EBP family member and NF-kappa B/p50. *J Exp Med* 1995;181:181–192.
580. Mikita T, Campbell D, Wu P, Williamson K, Schindler U. Requirements for interleukin-4-induced gene expression and functional characterization of Stat6. *Mol Cell Biol* 1996;16:5811–5820.
581. Shimoda K, van DJ, Sangster MY, et al. Lack of IL-4-induced Th2 response and IgE class switching in mice with disrupted Stat6 gene. *Nature* 1996;380: 630–633.
582. Takeda K, Tanaka T, Shi W, et al. Essential role of Stat6 in IL-4 signalling. *Nature* 1996;380:627–630.
583. Cooper C, Henderson A, Artandi S, Avitahl N, Calame K. Ig/EBP (C/EBP gamma) is a transdominant negative inhibitor of C/EBP family transcriptional activators. *Nucleic Acids Res* 1995;23:4371–4377.
584. Cooper CL, Berrier AL, Roman C, Calame KL. Limited expression of C/EBP family proteins during B lymphocyte development. Negative regulator Ig/EBP predominates early and activator NF-IL-6 is induced later. *J Immunol* 1994;153: 5049–5058.
585. Iciek LA, Delphin SA, Stavnezer J. CD40 cross-linking induces Ig epsilon germline transcripts in B cells via activation of NF-kappa B: Synergy with IL-4 induction. *J Immunol* 1997;158:4769–4779.
586. Sha WC, Liou HC, Tuomanen EI, Baltimore D. Targeted disruption of the p50 subunit of NF-kappa B leads to multifocal defects in immune responses. *Cell* 1995;80:321–330.
587. Kim J, Reeves R, Rothman P, Boothby M. The non-histone chromosomal protein HMG-I(Y) contributes to repression of the immunoglobulin heavy chain germline epsilon RNA promoter. *Eur J Immunol* 1995;25:798–808.
588. Wang DZ, Ray P, Boothby M. Interleukin 4-inducible phosphorylation of HMG-I(Y) is inhibited by rapamycin. *J Biol Chem* 1995;270:22924–22932.
589. Liao F, Birshtein BK, Busslinger M, Rothman P. The transcription factor BSAP (NF-HB) is essential for immunoglobulin germ-line epsilon transcription. *J Immunol* 1994;152:2904–2911.
589a. Thienes CP, De ML, Monticelli S, Busslinger M, Gould HJ, Vercelli D. The transcription factor B cell-specific activator protein (BSAP) enhances both IL-4- and CD40-mediated activation of the human epsilon germline promoter. *J Immunol* 1997;158:5874–5882.
590. Church GM, Ephrussi A, Gilbert W, Tonegawa S. Cell-type-specific contacts to immunoglobulin enhancers in nuclei. *Nature* 1985;313:798–801.
591. Ephrussi A, Church GM, Tonegawa S, Gilbert W. B lineage–specific interactions of an immunoglobulin enhancer with cellular factors in vivo. *Science* 1985;227: 134–140.
592. Kadesch T, Zervos P, Ruezinsky D. Functional analysis of the murine IgH enhancer: Evidence for negative control of cell-type specificity. *Nucleic Acids Res* 1986;14:8209–8221.
593. Murre C, McCaw PS, Baltimore D. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 1989;56:777–783.
594. Murre C, McCaw PS, Vaessin H, et al. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* 1989;58:537–544.
595. Lassar AB, Davis RL, Wright WE, et al. Functional activity of myogenic HLH proteins requires hetero-oligomerization with E12/E47-like proteins in vivo. *Cell* 1991;66:305–315.
596. Lin H, Yutzey KE, Konieczny SF. Muscle-specific expression of the troponin I gene requires interactions between helix-loop-helix muscle regulatory factors and ubiquitous transcription factors. *Mol Cell Biol* 1991;11:267–280.
597. Voliva CF, Aronheim A, Walker MD, Peterlin BM. B-cell factor 1 is required for optimal expression of the DRA promoter in B cells. *Mol Cell Biol* 1992;12: 2383–2390.
598. Shen CP, Kadesch T. B-cell-specific DNA binding by an E47 homodimer. *Mol Cell Biol* 1995;15:4518–4524.
599. Sun XH, Copeland NG, Jenkins NA, Baltimore D. Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. *Mol Cell Biol* 1991;11:5603–5611.
600. Wilson RB, Kiledjian M, Shen CP, et al. Repression of immunoglobulin enhancers by the helix-loop-helix protein Id: Implications for B-lymphoid-cell development. *Mol Cell Biol* 1991;11:6185–6191.
601. Bain G, Maandag EC, Izon DJ, et al. E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. *Cell* 1994;79:885–892.
602. Zhuang Y, Soriano P, Weintraub H. The helix-loop-helix gene E2A is required for B cell formation. *Cell* 1994;79:875–884.
603. Bain G, Robanus ME, te RH, et al. Both E12 and E47 allow commitment to the B cell lineage. *Immunity* 1997;6:145–154.
604. Schlissel M, Voronova A, Baltimore D. Helix-loop-helix transcription factor E47 activates germ-line immunoglobulin heavy-chain gene transcription and rearrangement in a pre-T-cell line. *Genes Dev* 1991;5:1367–1376.
605. Choi JK, Shen CP, Radomska HS, Eckhardt LA, Kadesch T. E47 activates the Ig-heavy chain and TdT loci in non-B cells. *EMBO J* 1996;15:5014–5021.
606. Sloan SR, Shen CP, McCarrick WR, Kadesch T. Phosphorylation of E47 as a potential determinant of B-cell-specific activity. *Mol Cell Biol* 1996;16: 6900–6908.
607. Bain G, Gruenwald S, Murre C. E2A and E2-2 are subunits of B-cell-specific E2-box DNA-binding proteins. *Mol Cell Biol* 1993;13:3522–3529.
608. Zhuang Y, Cheng P, Weintraub H. B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. *Mol Cell Biol* 1996;16:2898–2905.
609. Ruezinsky D, Beckmann H, Kadesch T. Modulation of the IgH enhancer's cell type specificity through a genetic switch. *Genes Dev* 1991;5:29–37.
610. Genetta T, Ruezinsky D, Kadesch T. Displacement of an E-box-binding repressor by basic helix-loop-helix proteins: Implications for B-cell specificity of the immunoglobulin heavy-chain enhancer. *Mol Cell Biol* 1994;14:6153–6163.
611. Carter RS, Ordentlich P, Kadesch T. Selective utilization of basic helix-loop-helix-leucine zipper proteins at the immunoglobulin heavy-chain enhancer. *Mol Cell Biol* 1997;17:18–23.
612. Merrell K, Wells S, Henderson A, et al. The absence of the transcription activator TFE3 impairs activation of B cells in vivo. *Mol Cell Biol* 1997;17: 3335–3344.
613. Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H. The protein Id: A negative regulator of helix-loop-helix DNA binding proteins. *Cell* 1990;61:49–59.
614. Sun XH. Constitutive expression of the Id1 gene impairs mouse B cell development. *Cell* 1994;79:893–900.
615. Mellentin JD, Murre C, Donlon TA, et al. The gene for enhancer binding proteins E12/E47 lies at the t(1;19) breakpoint in acute leukemias. *Science* 1989;246: 379–382.
616. Inaba T, Roberts WM, Shapiro LH, et al. Fusion of the leucine zipper gene HLF to the E2A gene in human acute B-lineage leukemia. *Science* 1992;257:531–534.
617. Nelsen B, Tian G, Erman B, et al. Regulation of lymphoid-specific immunoglobulin mu heavy chain gene enhancer by ETS-domain proteins. *Science* 1993;261: 82–86.
618. Libermann TA, Baltimore D. Pi, a pre-B-cell-specific enhancer element in the immunoglobulin heavy-chain enhancer. *Mol Cell Biol* 1993;13:5957–5969.
619. Akbarali Y, Oettgen P, Boltax J, Libermann TA. ELF-1 interacts with and transactivates the IgH enhancer pi site. *J Biol Chem* 1996;271:26007–26012.

620. Klemsz MJ, McKercher SR, Celada A, Van BC, Maki RA. The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. *Cell* 1990;61:113–124.
621. Scott EW, Simon MC, Anastasi J, Singh H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* 1994;265: 1573–1577.
622. Nikolajczyk BS, Nelsen B, Sen R. Precise alignment of sites required for mu enhancer activation in B cells. *Mol Cell Biol* 1996;16:4544–4554.
623. Nikolajczyk BS, Cortes M, Feinman R, Sen R. Combinatorial determinants of tissue-specific transcription in B cells and macrophages. *Mol Cell Biol* 1997;17: 3527–3535.
624. Rao E, Dang W, Tian G, Sen R. A three-protein-DNA complex on a B cell-specific domain of the immunoglobulin mu heavy chain gene enhancer. *J Biol Chem* 1997;272:6722–6732.
624a. Dang W, Sun XH, Sen R. ETS-mediated cooperation between basic helix-loop-helix motifs of the immunoglobulin mu heavy-chain gene enhancer. *Mol Cell Biol* 1998;18:1477-1488.
624b. Peterson CL, Eaton S, Calame K. Purified mu EBP-E binds to immunoglobulin enhancers and promoters. *Mol Cell Biol* 1988;8:4972-4980.
624c. Tsao BP, Wang XF, Peterson CL, Calame K. In vivo functional analysis of in vitro protein binding sites in the immunoglobulin heavy chain enhancer. *Nucleic Acids Res* 1988;16:3239-3253.
625. Raynal MC, Liu ZY, Hirano T, Mayer L, Kishimoto T, Chen KS. Interleukin 6 induces secretion of IgG1 by coordinated transcriptional activation and differential mRNA accumulation. *Proc Natl Acad Sci USA* 1989;86:8024–8028.
626. Park K, Atchison ML. Isolation of a candidate repressor/activator, NF-E1 (YY-1, delta), that binds to the immunoglobulin kappa 3' enhancer and the immunoglobulin heavy-chain mu E1 site. *Proc Natl Acad Sci USA* 1991;88:9804–9808.
627. Shi Y, Seto E, Chang LS, Shenk T. Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein. *Cell* 1991;67:377–388.
628. Bushmeyer S, Park K, Atchison ML. Characterization of functional domains within the multifunctional transcription factor, YY1. *J Biol Chem* 1995;270: 30213–30220.
629. Kiledjian M, Su LK, Kadesch T. Identification and characterization of two functional domains within the murine heavy-chain enhancer. *Mol Cell Biol* 1988;8: 145–152.
630. Wasylyk C, Wasylyk B. The immunoglobulin heavy-chain B-lymphocyte enhancer efficiently stimulates transcription in non-lymphoid cells. *EMBO J* 1986;5:553–560.
631. Imler JL, Lemaire C, Wasylyk C, Wasylyk B. Negative regulation contributes to tissue specificity of the immunoglobulin heavy-chain enhancer. *Mol Cell Biol* 1987;7:2558–2567.
632. Weinberger J, Jat PS, Sharp PA. Localization of a repressive sequence contributing to B-cell specificity in the immunoglobulin heavy-chain enhancer. *Mol Cell Biol* 1988;8:988–992.
633. Cockerill PN, Yuen MH, Garrard WT. The enhancer of the immunoglobulin heavy chain locus is flanked by presumptive chromosomal loop anchorage elements. *J Biol Chem* 1987;262:5394–5397.
634. Scheuermann RH, Chen U. A developmental-specific factor binds to suppressor sites flanking the immunoglobulin heavy-chain enhancer. *Genes Dev* 1989;3: 1255–1266.
635. Zong RT, Scheuermann RH. Mutually exclusive interaction of a novel matrix attachment region binding protein and the NF-muNR enhancer repressor. Implications for regulation of immunoglobulin heavy chain expression. *J Biol Chem* 1995;270:24010–24018.
636. Forrester WC, van Genderen C, Jenuwein T, Grosschedl R. Dependence of enhancer-mediated transcription of the immunoglobulin mu gene on nuclear matrix attachment regions. *Science* 1994;265:1221–1225.
637. McKnight RA, Shamay A, Sankaran L, Wall RJ, Hennighausen L. Matrix-attachment regions can impart position-independent regulation of a tissue-specific gene in transgenic mice. *Proc Natl Acad Sci USA* 1992;89:6943–6947.
638. Jenuwein T, Forrester WC, Qiu RG, Grosschedl R. The immunoglobulin mu enhancer core establishes local factor access in nuclear chromatin independent of transcriptional stimulation. *Genes Dev* 1993;7:2016–2032.
639. Jenuwein T, Forrester WC, Fernandez HL, Laible G, Dull M, Grosschedl R. Extension of chromatin accessibility by nuclear matrix attachment regions. *Nature* 1997;385:269–272.
640. Dickinson LA, Joh T, Kohwi Y, Kohwi ST. A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell* 1992;70:631–645.
641. Dickinson LA, Kohwi ST. Nucleolin is a matrix attachment region DNA-binding protein that specifically recognizes a region with high base-unpairing potential. *Mol Cell Biol* 1995;15:456–465.
642. Klein S, Sablitzky F, Radbruch A. Deletion of the IgH enhancer does not reduce immunoglobulin heavy chain production of a hybridoma IgD class switch variant. *EMBO J* 1984;3:2473–2476.
643. Wabl MR, Burrows PD. Expression of immunoglobulin heavy chain at a high level in the absence of a proposed immunoglobulin enhancer element in cis. *Proc Natl Acad Sci USA* 1984;81:2452–2455.
644. Aguilera RJ, Hope TJ, Sakano H. Characterization of immunoglobulin enhancer deletions in murine plasmacytomas. *EMBO J* 1985;4:3689–3693.
644a. Klein U, Klein G, Ehlin HB, Rajewsky K, Kuppers R. Burkitt's lymphoma is a malignancy of mature B cells expressing somatically mutated V region genes. *Mol Med* 1995;1:495-505.
644b. Tamaru J, Hummel M, Marafioti T, et al. Burkitt's lymphomas express VH genes with a moderate number of antigen-selected somatic mutations. *Am J Pathol* 1995;147:1398-1407.
645. Eckhardt LA, Birshtein BK. Independent immunoglobulin class-switch events occurring in a single myeloma cell line. *Mol Cell Biol* 1985;5:856–868.
646. Gregor PD, Morrison SL. Myeloma mutant with a novel 3' flanking region: Loss of normal sequence and insertion of repetitive elements leads to decreased transcription but normal processing of the alpha heavy-chain gene products. *Mol Cell Biol* 1986;6:1903–1916.
647. Pettersson S, Cook GP, Bruggemann M, Williams GT, Neuberger MS. A second B cell-specific enhancer 3' of the immunoglobulin heavy-chain locus. *Nature* 1990;344:165–168.
648. Dariavach P, Williams GT, Campbell K, Pettersson S, Neuberger MS. The mouse IgH 3'-enhancer. *Eur J Immunol* 1991;21:1499–1504.
649. Lieberson R, Giannini SL, Birshtein BK, Eckhardt LA. An enhancer at the 3' end of the mouse immunoglobulin heavy chain locus. *Nucleic Acids Res* 1991;19:933–937.
650. Matthias P, Baltimore D. The immunoglobulin heavy chain locus contains another B-cell-specific 3' enhancer close to the alpha constant region. *Mol Cell Biol* 1993;13:1547–1553.
651. Chauveau C, Cogne M. Palindromic structure of the IgH 3'locus control region. *Nat Genet* 1996;14:15–16.
652. Saleque S, Singh M, Little RD, Giannini SL, Michaelson JS, Birshtein BK. Dyad symmetry within the mouse 3' IgH regulatory region includes two virtually identical enhancers (C alpha3'E and hs3). *J Immunol* 1997;158:4780–4787.
653. Meyer KB, Skogberg M, Margenfeld C, Ireland J, Pettersson S. Repression of the immunoglobulin heavy chain 3' enhancer by helix-loop-helix protein Id3 via a functionally important E47/E12 binding site: Implications for developmental control of enhancer function. *Eur J Immunol* 1995;25:1770–1777.
654. Neurath MF, Max EE, Strober W. Pax5 (BSAP) regulates the murine immunoglobulin 3' alpha enhancer by suppressing binding of NF-alpha P, a protein that controls heavy chain transcription. *Proc Natl Acad Sci USA* 1995;92: 5336–5340.
655. Ernst P, Smale ST. Combinatorial regulation of transcription II: the immunoglobulin mu heavy chain gene. *Immunity* 1995;2:427–438.
656. Singh M, Birshtein BK. Concerted repression of an immunoglobulin heavy-chain enhancer, 3' alpha E(hs1,2). *Proc Natl Acad Sci USA* 1996;93:4392–4397.
657. Linderson Y, Cross D, Neurath MF, Pettersson S. NFE, a new transcriptional activator that facilitates p50 and c-Rel-dependent IgH 3' enhancer activity. *Eur J Immunol* 1997;27:468–475.
658. Grant PA, Thompson CB, Pettersson S. IgM receptor-mediated transactivation of the IgH 3' enhancer couples a novel Elf-1-AP-1 protein complex to the developmental control of enhancer function. *EMBO J* 1995;14:4501–4513.
659. Grant PA, Andersson T, Neurath MF, et al. A T cell controlled molecular pathway regulating the IgH locus: CD40-mediated activation of the IgH 3' enhancer. *EMBO J* 1996;15:6691–6700.
660. Michaelson JS, Giannini SL, Birshtein BK. Identification of 3' alpha-hs4, a novel Ig heavy chain enhancer element regulated at multiple stages of B cell differentiation. *Nucleic Acids Res* 1995;23:975–981.
661. Michaelson JS, Singh M, Snapper CM, Sha WC, Baltimore D, Birshtein BK. Regulation of 3' IgH enhancers by a common set of factors, including kappa B-binding proteins. *J Immunol* 1996;156:2828–2839.
662. Chen J, Young F, Bottaro A, Stewart V, Smith RK, Alt FW. Mutations of the intronic IgH enhancer and its flanking sequences differentially affect accessibility of the JH locus. *EMBO J* 1993;12:4635–4645.
663. Serwe M, Sablitzky F. V(D)J recombination in B cells is impaired but not blocked by targeted deletion of the immunoglobulin heavy chain intron enhancer. *EMBO J* 1993;12:2321–2327.
664. Lieberson R, Ong J, Shi X, Eckhardt LA. Immunoglobulin gene transcription ceases upon deletion of a distant enhancer. *EMBO J* 1995;14:6229–6238.
665. Singh M, Birshtein BK. NF-HB (BSAP) is a repressor of the murine immunoglobulin heavy-chain 3' alpha enhancer at early stages of B-cell differentiation. *Mol Cell Biol* 1993;13:3611–3622.
666. Arulampalam V, Grant PA, Samuelsson A, Lendahl U, Pettersson S. Lipopolysaccharide-dependent transactivation of the temporally regulated immunoglobulin heavy chain 3'enhancer. *Eur J Immunol* 1994;24:1671–1677.
667. Arulampalam V, Furebring C, Samuelsson A, et al. Elevated expression levels of an Ig transgene in mice links the IgH 3' enhancer to the regulation of IgH expression. *Int Immunol* 1996;8:1149–1157.
668. Picard D, Schaffner W. A lymphocyte-specific enhancer in the mouse immunoglobulin kappa gene. *Nature* 1984;307:80–82.
669. Queen C, Stafford J. Fine mapping of an immunoglobulin gene activator. *Mol Cell Biol* 1984;4:1042–1049.
670. Chung SY, Folsom V, Wooley J. DNase I-hypersensitive sites in the chromatin of immunoglobulin kappa light chain genes. *Proc Natl Acad Sci USA* 1983;80: 2427–2431.
671. Emorine L, Kuehl M, Weir L, Leder P, Max EE. A conserved sequence in the immunoglobulin J kappa-C kappa intron: Possible enhancer element. *Nature* 1983;304:447–449.
672. Sen R, Baltimore D. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* 1986;46:705–716.
673. Sen R, Baltimore D. Inducibility of kappa immunoglobulin enhancer-binding protein Nf-kappa B by a posttranslational mechanism. *Cell* 1986;47:921–928.
674. Pierce JW, Lenardo M, Baltimore D. Oligonucleotide that binds nuclear factor

NF-kappa B acts as a lymphoid-specific and inducible enhancer element. *Proc Natl Acad Sci USA* 1988;85:1482–1486.
675. Rooney JW, Dubois PM, Sibley CH. Cross-linking of surface IgM activates NF-kappa B in B lymphocyte. *Eur J Immunol* 1991;21:2993–2998.
676. Baeuerle PA, Baltimore D. Activation of DNA-binding activity in an apparently cytoplasmic precursor of the NF-kappa B transcription factor. *Cell* 1988;53: 211–217.
677. Baeuerle PA, Baltimore D. I kappa B: A specific inhibitor of the NF-kappa B transcription factor. *Science* 1988;242:540–546.
678. Baldwin AS. The NF-kB and IkB proteins: New discoveries and insights. *Annu Rev Immunol* 1996;14:649–681.
679. Stephens RM, Rice NR, Hiebsch RR, Bose HJ, Gilden RV. Nucleotide sequence of v-rel: The oncogene of reticuloendotheliosis virus. *Proc Natl Acad Sci USA* 1983;80:6229–6233.
680. Lernbecher T, Muller U, Wirth T. Distinct NF-kappa B/Rel transcription factors are responsible for tissue-specific and inducible gene activation. *Nature* 1993;365:767–770.
681. Liou HC, Sha WC, Scott ML, Baltimore D. Sequential induction of NF-kappa B/Rel family proteins during B-cell terminal differentiation. *Mol Cell Biol* 1994;14:5349–5359.
682. Miyamoto S, Schmitt MJ, Verma IM. Qualitative changes in the subunit composition of kappa B-binding complexes during murine B-cell differentiation. *Proc Natl Acad Sci USA* 1994;91:5056–5060.
683. Davis N, Ghosh S, Simmons DL, et al. Rel-associated pp40: An inhibitor of the rel family of transcription factors. *Science* 1991;253:1268–1271.
684. Haskill S, Beg AA, Tompkins SM, et al. Characterization of an immediate-early gene induced in adherent monocytes that encodes I kappa B-like activity. *Cell* 1991;65:1281–1289.
685. Thompson JE, Phillips RJ, Erdjument BH, Tempst P, Ghosh S. I kappa B-beta regulates the persistent response in a biphasic activation of NF-kappa B. *Cell* 1995;80:573–582.
686. Zhang Q, Didonato JA, Karin M, McKeithan TW. BCL3 encodes a nuclear protein which can alter the subcellular location of NF-kappa B proteins. *Mol Cell Biol* 1994;14:3915–3926.
687. Bours V, Franzoso G, Azarenko V, et al. The oncoprotein Bcl-3 directly transactivates through kappa B motifs via association with DNA-binding p50B homodimers. *Cell* 1993;72:729–739.
688. Fujita T, Nolan GP, Liou HC, Scott ML, Baltimore D. The candidate proto-oncogene bcl-3 encodes a transcriptional coactivator that activates through NF-kappa B p50 homodimers. *Genes Dev* 1993;7:1354–1363.
689. Naumann M, Scheidereit C. Activation of NF-kappa B in vivo is regulated by multiple phosphorylations. *EMBO J* 1994;13:4597–4607.
690. Chen JC, Parent L, Maniatis T. Site-specific phosphorylation of I-kappa B-alpha by a novel ubiquitination-dependent protein kinase activity. *Cell* 1996;84: 853–862.
691. DiDonato JA, Hayakawa M, Rothwarf DM, Zandi E, Karin M. A cytokine-responsive I-kappaB kinase that activates the transcription factor NF-kappaB. *Nature* 1997;388:548–554.
691a. Regnier CH, Song HY, Gao X, Goeddel DV, Cao Z, Rothe M. Identification and characterization of an I-kappa-B kinase. *Cell* 1997;90:373-383.
692. Chen Z, Hagler J, Palombella VJ, et al. Signal-induced site-specific phosphorylation targets I kappa B alpha to the ubiquitin-proteasome pathway. *Genes Dev* 1995;9:1586–1597.
693. Arenzana SF, Thompson J, Rodriguez MS, Bachelerie F, Thomas D, Hay RT. Inducible nuclear expression of newly synthesized I kappa B alpha negatively regulates DNA-binding and transcriptional activities of NF-kappa B. *Mol Cell Biol* 1995;15:2689–2696.
694. Scherer DC, Brockman JA, Bendall HH, Zhang GM, Ballard DW, Oltz EM. Corepression of RelA and c-rel inhibits immunoglobulin kappa gene transcription and rearrangement in precursor B lymphocytes. *Immunity* 1996;5:563–574.
695. Cockerill PN, Garrard WT. Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *Cell* 1986;44:273–282.
696. Blasquez VC, Xu M, Moses SC, Garrard WT. Immunoglobulin kappa gene expression after stable integration. I. Role of the intronic MAR and enhancer in plasmacytoma cells. *J Biol Chem* 1989;264:21183–21189.
697. Xu M, Hammer RE, Blasquez VC, Jones SL, Garrard WT. Immunoglobulin kappa gene expression after stable integration. II. Role of the intronic MAR and enhancer in transgenic mice. *J Biol Chem* 1989;264:21190–21195.
698. Lichtenstein M, Keini G, Cedar H, Bergman Y. B cell-specific demethylation: A novel role for the intronic kappa chain enhancer sequence. *Cell* 1994;76: 913–923.
699. Kirillov A, Kistler B, Mostoslavsky R, Cedar H, Wirth T, Bergman Y. A role for nuclear NF-kappaB in B-cell-specific demethylation of the Igkappa locus. *Nat Genet* 1996;13:435–441.
700. Schanke JT, Marcuzzi A, Podzorski RP, Van NB. An AP1 binding site upstream of the kappa immunoglobulin intron enhancer binds inducible factors and contributes to expression. *Nucleic Acids Res* 1994;22:5425–5432.
701. Pierce JW, Gifford AM, Baltimore D. Silencing of the expression of the immunoglobulin kappa gene in non-B cells. *Mol Cell Biol* 1991;11:1431–1437.
702. Saksela K, Baltimore D. Negative regulation of immunoglobulin kappa light-chain gene transcription by a short sequence homologous to the murine B1 repetitive element. *Mol Cell Biol* 1993;13:3698–3705.
703. Atchison ML, Perry RP. The role of the kappa enhancer and its binding factor NF-kappa B in the developmental regulation of kappa gene transcription. *Cell* 1987;48:121–128.
704. Meyer KB, Neuberger MS. The immunoglobulin kappa locus contains a second, stronger B-cell-specific enhancer which is located downstream of the constant region. *EMBO J* 1989;8:1959–1964.
705. Meyer KB, Sharpe MJ, Surani MA, Neuberger MS. The importance of the 3'-enhancer region in immunoglobulin kappa gene expression. *Nucleic Acids Res* 1990;18:5609–5615.
706. Pongubala JM, Atchison ML. Functional characterization of the developmentally controlled immunoglobulin kappa 3' enhancer: Regulation by Id, a repressor of helix-loop-helix transcription factors. *Mol Cell Biol* 1991;11:1040–1047.
707. Pongubala JM, Nagulapalli S, Klemsz MJ, et al. PU.1 recruits a second nuclear factor to a site important for immunoglobulin kappa 3' enhancer activity. *Mol Cell Biol* 1992;12:368–378.
708. Brass AL, Kehrli E, Eisenbeis CF, Storb U, Singh H. Pip, a lymphoid-restricted IRF, contains a regulatory domain that is important for autoinhibition and ternary complex formation with the Ets factor PU.1. *Genes Dev* 1996;10: 2335–2347.
709. Pongubala JM, Atchison ML. Activating transcription factor 1 and cyclic AMP response element modulator can modulate the activity of the immunoglobulin kappa 3' enhancer. *J Biol Chem* 1995;270:10304–10313.
710. Pongubala JM, Atchison ML. PU.1 can participate in an active enhancer complex without its transcriptional activation domain. *Proc Natl Acad Sci USA* 1997;94: 127–132.
711. Roque MC, Smith PA, Blasquez VC. A developmentally modulated chromatin structure at the mouse immunoglobulin kappa 3' enhancer. *Mol Cell Biol* 1996;16:3138–3155.
712. Costa MW, Atchison ML. Identification of an Spl-like element within the immunoglobulin kappa 3' enhancer necessary for maximal enhancer activity. *Biochemistry* 1996;35:8662–8669.
713. Meyer KB, Ireland J. Activation of the immunoglobulin kappa 3' enhancer in pre-B cells correlates with the suppression of a nuclear factor binding to a sequence flanking the active core. *Nucleic Acids Res* 1994;22:1576–1582.
714. Takeda S, Zou YR, Bluethmann H, Kitamura D, Muller U, Rajewsky K. Deletion of the immunoglobulin kappa chain intron enhancer abolishes kappa chain gene rearrangement in cis but not lambda chain gene rearrangement in trans. *EMBO J* 1993;12:2329–2336.
715. Fulton R, Van Ness B. Kappa immunoglobulin promoters and enhancers display developmentally controlled interactions. *Nucleic Acids Res* 1993;21:4941–4947.
716. Fulton R, van NB. Selective synergy of immunoglobulin enhancer elements in B-cell development: A characteristic of kappa light chain enhancers, but not heavy chain enhancers. *Nucleic Acids Res* 1994;22:4216–4223.
717. Meyer KB, Teh YM, Neuberger MS. The Ig kappa 3'-enhancer triggers gene expression in early B lymphocytes but its activity is enhanced on B cell activation. *Int Immunol* 1996;8:1561–1568.
718. Shaffer AL, Peng A, Schlissel MS. In vivo occupancy of the kappa light chain enhancers in primary pro- and pre-B cells: A model for kappa locus activation. *Immunity* 1997;6:131–143.
719. Hagman J, Rudin CM, Haasch D, Chaplin D, Storb U. A novel enhancer in the immunoglobulin lambda locus is duplicated and functionally independent of NF kappa B. *Genes Dev* 1990;4:978–992.
720. Rudin CM, Storb U. Two conserved essential motifs of the murine immunoglobulin lambda enhancers bind B-cell-specific factors. *Mol Cell Biol* 1992;12: 309–320.
721. Eisenbeis CF, Singh H, Storb U. PU.1 is a component of a multiprotein complex which binds an essential site in the murine immunoglobulin lambda 2-4 enhancer. *Mol Cell Biol* 1993;13:6452–6461.
722. Eisenbeis CF, Singh H, Storb U. Pip, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcriptional activator. *Genes Dev* 1995;9: 1377–1387.
723. Blomberg BB, Rudin CM, Storb U. Identification and localization of an enhancer for the human lambda L chain Ig gene complex. *J Immunol* 1991;147: 2354–2358.
724. Asenbauer H, Klobeck HG. Tissue-specific deoxyribonuclease I-hypersensitive sites in the vicinity of the immunoglobulin C lambda cluster of man. *Eur J Immunol* 1996;26:142–150.
725. Glozak MA, Blomberg BB. The human lambda immunoglobulin enhancer is controlled by both positive elements and developmentally regulated negative elements. *Mol Immunol* 1996;33:427–438.
726. Chen C, Birshtein BK. Virtually identical enhancers containing a segment of homology to murine 3'IgH-E(hs1,2) lie downstream of human Ig C alpha 1 and C alpha 2 genes. *J Immunol* 1997;159:1310–1318.
727. Pospelov VA, Klobeck HG, Zachau HG. Correlation between DNase I hypersensitive sites and putative regulatory sequences in human immunoglobulin genes of the kappa light chain type. *Nucleic Acids Res* 1984;12:7007–7021.
728. Potter H, Weir L, Leder P. Enhancer-dependent expression of human kappa immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. *Proc Natl Acad Sci USA* 1984;81:7161–7165.
729. Gimble JM, Levens D, Max EE. B-cell nuclear proteins binding in vitro to the human immunoglobulin kappa enhancer: Localization by exonuclease protection. *Mol Cell Biol* 1987;7:1815–1822.
730. Judde JG, Max EE. Characterization of the human immunoglobulin kappa gene 3' enhancer: Functional importance of three motifs that demonstrate B-cell-specific in vivo footprints. *Mol Cell Biol* 1992;12:5206–5216.

731. Rabbitts TH, Forster A, Baer R, Hamlyn PH. Transcription enhancer identified near the human C mu immunoglobulin heavy chain gene is unavailable to the translocated c-myc gene in a Burkitt lymphoma. *Nature* 1983;306:806–809.
732. Hayday AC, Gillies SD, Saito H, et al. Activation of a translocated human c-myc gene by an enhancer in the immunoglobulin heavy-chain locus. *Nature* 1984; 307:334–340.
733. Okamura K, Ishiguro H, Ichihara Y, Kurosawa Y. Comparison of nucleotide sequences from upstream of the DQ52 gene to the S mu region of immunoglobulin heavy-chain gene loci between Suncus murinus, mouse and human. *Mol Immunol* 1993;30:461–467.
734. Enjoji M. Human HE2 (microB) and microA motifs show the same function as whole IgH intronic enhancer in transgenic mice. *Mol Cell Biochem* 1994;137: 33–37.
735. Waldmann TA. The arrangement of immunoglobulin and T cell receptor genes in human lymphoproliferative disorders. *Adv Immunol* 1987;40:247–321.
736. Korsmeyer SJ. B-lymphoid neoplasms: immunoglobulin genes as molecular determinants of clonality, lineage, differentiation, and translocation. *Adv Intern Med* 1988;33:1–15.
737. Felix CA, Poplack DG. Characterization of acute lymphoblastic leukemia of childhood by immunoglobulin and T-cell receptor gene patterns. *Leukemia* 1991; 5:1015–1025.
738. Veronese ML, Schichman SA, Croce CM. Molecular diagnosis of lymphoma. *Curr Opin Oncol* 1996;8:346–352.
739. Roberts WM, Estrov Z, Kitchingman GR, Zipf TF. The clinical significance of residual disease in childhood acute lymphoblastic leukemia as detected by polymerase chain reaction amplification by antigen-receptor gene sequences. *Leuk Lymphoma* 1996;20:181–197.
740. Felix CA, Wright JJ, Poplack DG, et al. T cell receptor alpha-, beta-, and gamma-genes in T cell and pre-B cell acute lymphoblastic leukemia. *J Clin Invest* 1987;80:545–556.
741. Wright JJ, Poplack DG, Bakhshi A, et al. Gene rearrangements as markers of clonal variation and minimal residual disease in acute lymphoblastic leukemia. *J Clin Oncol* 1987;5:735–741.
742. Croce CM, Shander M, Martinis J, et al. Chromosomal location of the genes for human immunoglobulin heavy chains. *Proc Natl Acad Sci USA* 1979;76: 3416–3419.
743. Taub R, Kelly K, Battey J, et al. A novel alteration in the structure of an activated c-myc gene in a variant t(2;8) Burkitt lymphoma. *Cell* 1984;37:511–520.
744. Showe LC, Croce CM. The role of chromosomal translocations in B- and T-cell neoplasia. *Ann Rev Immunol* 1987;5:253–277.
745. Kato GJ, Lee WM, Chen LL, Dang CV. Max: Functional domains and interaction with c-Myc. *Genes Dev* 1992;6:81–92.
746. Galaktionov K, Chen X, Beach D. Cdc25 cell-cycle phosphatase as a target of c-myc. *Nature* 1996;382:511–517.
747. Schuldiner O, Eden A, Ben YT, Yanuka O, Simchen G, Benvenisty N. ECA39, a conserved gene regulated by c-Myc in mice, is involved in G1/S cell cycle regulation in yeast. *Proc Natl Acad Sci USA* 1996;93:7143–7148.
748. Nishikura K, ar-Rushdi A, Erikson J, Watt R, Rovera G, Croce CM. Differential expression of the normal and of the translocated human c-myc oncogenes in B cells. *Proc Natl Acad Sci USA* 1983;80:4822–4826.
749. Shima EA, Le BM, McKeithan TW, et al. Gene encoding the alpha chain of the T-cell receptor is moved immediately downstream of c-myc in a chromosomal 8;14 translocation in a cell line from a human T-cell leukemia. *Proc Natl Acad Sci USA* 1986;83:3439–3443.
750. Tsujimoto Y, Yunis J, Onorato SL, Erikson J, Nowell PC, Croce CM. Molecular cloning of the chromosomal breakpoint of B-cell lymphomas and leukemias with the t(11;14) chromosome translocation. *Science* 1984;224: 1403–1406.
751. Rosenberg CL, Wong E, Petty EM, et al. PRAD1, a candidate BCL1 oncogene: Mapping and expression in centrocytic lymphoma. *Proc Natl Acad Sci USA* 1991;88:9638–9642.
752. de Boer C, van Krieken J, Schuuring E, Kluin PM. Bcl-1/cyclin D1 in malignant lymphoma. *Ann Oncol* 1997;2:109–117.
753. Tsujimoto Y, Finger LR, Yunis J, Nowell PC, Croce CM. Cloning of the chromosome breakpoint of neoplastic B cells with the t(14;18) chromosome translocation. *Science* 1984;226:1097–1099.
754. Hockenbery D, Nuñez G, Milliman C, Schreiber RD, Korsmeyer SJ. Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. *Nature* 1990;348:334–336.
755. Reed JC. Regulation of apoptosis by bcl-2 family proteins and its role in cancer and chemoresistance. *Curr Opin Oncol* 1995;7:541–546.
756. Hockenbery DM, Zutter M, Hickey W, Nahm M, Korsmeyer SJ. BCL2 protein is topographically restricted in tissues characterized by apoptotic cell death. *Proc Natl Acad Sci USA* 1991;88:6961–6965.
757. McDonnell TJ, Deane N, Platt FM, et al. bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation. *Cell* 1989;57:79–88.
758. McKeithan TW, Rowley JD, Shows TB, Diaz MO. Cloning of the chromosome translocation breakpoint junction of the t(14;19) in chronic lymphocytic leukemia. *Proc Natl Acad Sci USA* 1987;84:9257–9260.
759. Michaux L, Dierlamm J, Wlodarska I, et al. t(14;19)/BCL3 rearrangements in lymphoproliferative disorders: A review of 23 cases. *Cancer Genet Cytogenet* 1997;94:36–43.
760. Ohno H, Takimoto G, McKeithan TW. The candidate proto-oncogene bcl-3 is related to genes implicated in cell lineage determination and cell cycle control. *Cell* 1990;60:991–997.
761. Baron BW, Nucifora G, McCabe N, Espinosa Rd, Le BM, McKeithan TW. Identification of the gene associated with the recurring chromosomal translocations t(3;14)(q27;q32) and t(3;22)(q27;q11) in B-cell lymphomas. *Proc Natl Acad Sci USA* 1993;90:5262–5266.
762. Seyfert VL, Allman D, He Y, Staudt LM. Transcriptional repression by the proto-oncogene BCL-6. *Oncogene* 1996;12:2331–2342.
763. Allman D, Jain A, Dent A, et al. BCL-6 expression during B-cell activation. *Blood* 1996;87:5257–5268.
764. Dent AL, Shaffer AL, Yu X, Allman D, Staudt LM. Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science* 1997; 276:589–592.
765. Ye BH, Cattoretti G, Shen Q, et al. The BCL-6 proto-oncogene controls germinal-centre formation and Th2- type inflammation. *Nat Genet* 1997;16:161–170.
766. Zani VJ, Asou N, Jadayel D, et al. Molecular cloning of complex chromosomal translocation t(8;14;12)(q24.1;q32.3;q24.1) in a Burkitt lymphoma cell line defines a new gene (BCL7A) with homology to caldesmon. *Blood* 1996;87: 3124–3134.
767. Dyomin VG, Rao PH, Dalla-Favera R, Chaganti R. BCL8, a novel gene involved in translocations affecting band 15q11-13 in diffuse large-cell lymphoma. *Proc Natl Acad Sci USA* 1997;94:5728–5732.
768. Akasaka T, Muramatsu M, Ohno H, et al. Application of long-distance polymerase chain reaction to detection of junctional sequences created by chromosomal translocation in mature B-cell neoplasms. *Blood* 1996;88:985–994.
769. Denny CT, Yoshikai Y, Mak TW, Smith SD, Hollis GF, Kirsch IR. A chromosome 14 inversion in a T-cell lymphoma is caused by site-specific recombination between immunoglobulin and T-cell receptor loci. *Nature* 1986;320: 549–551.
770. Baer R, Forster A, Rabbitts TH. The mechanism of chromosome 14 inversion in a human T cell lymphoma. *Cell* 1987;50:97–105.
771. Lipkowitz S, Garry VF, Kirsch IR. Interlocus V-J recombination measures genomic instability in agriculture workers at risk for lymphoid malignancies. *Proc Natl Acad Sci USA* 1992;89:5301–5305.
772. Stavnezer-Nordgren J, Kekish O, Zegers BJ. Molecular defects in a human immunoglobulin kappa chain deficiency. *Science* 1985;230:458–461.
772a. Yel L, Minegishi Y, Coustan SE, et al. Mutations in the mu heavy-chain gene in patients with agammaglobulinemia. *N Engl J Med* 1996;335:1486-1493.
773. Humphries CG, Shen A, Kuziel WA, Capra JD, Blattner FR, Tucker PW. A new human immunoglobulin VH family preferentially rearranged in immature B-cell tumours. *Nature* 1988;331:446–449.
774. Chen PP, Siminovitch KA, Olsen NJ, Erger RA, Carson DA. A highly informative probe for two polymorphic Vh gene regions that contain one or more autoantibody-associated Vh genes. *J Clin Invest* 1989;84:706–710.
775. Sanz I, Kelly P, Williams C, Scholl S, Tucker P, Capra JD. The smaller human VH gene families display remarkably little polymorphism. *EMBO J* 1989;8: 3741–3748.
776. Shin EK, Matsuda F, Nagaoka H, et al. Physical map of the 3' region of the human immunoglobulin heavy chain locus: Clustering of autoantibody-related variable segments in one haplotype. *EMBO J* 1991;10:3641–3645.
777. Walter MA, Gibson WT, Ebers GC, Cox DW. Susceptibility to multiple sclerosis is associated with the proximal immunoglobulin heavy chain variable region. *J Clin Invest* 1991;87:1266–1273.
778. Moxley G. DNA polymorphism of immunoglobulin kappa confers risk of rheumatoid arthritis. *Arthritis Rheum* 1989;32:634–637.
779. Meindl A, Klobeck HG, Ohnheiser R, Zachau HG. The V kappa gene repertoire in the human germ line. *Eur J Immunol* 1990;20:1855–1863.
780. Feeney AJ, Atkinson MJ, Cowan MJ, Escuro G, Lugo G. A defective Vkappa A2 allele in Navajos which may play a role in increased susceptibility to haemophilus influenzae type b disease. *J Clin Invest* 1996;97:2277–2282.
781. Baker MD, Wu GE, Toone WM, Murialdo H, Davis AC, Shulman MJ. A region of the immunoglobulin-mu heavy chain necessary for forming pentameric IgM. *J Immunol* 1986;137:1724–1728.
782. Shulman MJ, Collins C, Pennell N, Hozumi N. Complement activation by IgM: Evidence for the importance of the third constant domain of the mu heavy chain. *Eur J Immunol* 1987;17:549–554.
783. Tao MH, Canfield SM, Morrison SL. The differential ability of human IgG1 and IgG4 to activate complement is determined by the COOH-terminal sequence of the CH2 domain. *J Exp Med* 1991;173:1025–1028.
784. Duncan AR, Winter G. The binding site for C1q on IgG. *Nature* 1988;332: 738–740.
785. Helm B, Marsh P, Vercelli D, Padlan E, Gould H, Geha R. The mast cell binding site on human immunoglobulin E. *Nature* 1988;331:180–183.
786. Roberts S, Cheetham JC, Rees AR. Generation of an antibody with enhanced affinity and specificity for its antigen by protein engineering. *Nature* 1987;328: 731–734.
787. Boulianne GL, Hozumi N, Shulman MJ. Production of functional chimaeric mouse/human antibody. *Nature* 1984;312:643–646.
788. Morrison SL. Transfectomas provide novel chimeric antibodies. *Science* 1985; 229:1202–1207.
789. Liu AY, Robinson RR, Murray EJ, Ledbetter JA, Hellstrom I, Hellstrom KE. Production of a mouse-human chimeric monoclonal antibody to CD20 with potent Fc-dependent biologic activity. *J Immunol* 1987;139:3521–3526.
790. Jones PT, Dear PH, Foote J, Neuberger MS, Winter G. Replacing the comple-

mentarity-determining regions in a human antibody with those from a mouse. *Nature* 1986;321:522–525.
791. Co MS, Avdalovic NM, Caron PC, Avdalovic MV, Scheinberg DA, Queen C. Chimeric and humanized antibodies with specificity for the CD33 antigen. *J Immunol* 1992;148:1149–1154.
792. Mendez MJ, Green LL, Corvalan JR, et al. Functional transplant of megabase human immunoglobulin loci recapitulates human antibody response in mice. *Nat Genet* 1997;15:146–156.
793. Desmyter A, Transue TR, Ghahroudi MA, et al. Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat Struct Biol* 1996;3:803–811.
794. Bird RE, Hardman KD, Jacobson JW, et al. Single-chain antigen-binding proteins. *Science* 1988;242:423–426.
795. Huston JS, Levinson D, Mudgett HM, et al. Protein engineering of antibody binding sites: Recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. *Proc Natl Acad Sci USA* 1988;85: 5879–5883.
796. Clackson T, Hoogenboom HR, Griffiths AD, Winter G. Making antibody fragments using phage display libraries. *Nature* 1991;352:624–628.
797. Kang AS, Barbas CF, Janda KD, Benkovic SJ, Lerner RA. Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proc Natl Acad Sci USA* 1991;88:4363–4366.
798. Huse WD, Stinchcombe TJ, Glaser SM, et al. Application of a filamentous phage pVIII fusion protein system suitable for efficient production, screening and mutagenesis of F(ab) antibody fragments. *J Immunol* 1992;149: 3914–3920.
799. Yelton DE, Rosok MJ, Cruz G, et al. Affinity maturation of the BR96 anti-carcinoma antibody by codon-based mutagenesis. *J Immunol* 1995;155:1994–2004.
800. Crameri A, Cwirla S, Stemmer WP. Construction and evolution of antibody-phage libraries by DNA shuffling. *Nat Med* 1996;2:100–102.
801. Orlandi R, Gussow DH, Jones PT, Winter G. Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. *Proc Natl Acad Sci USA* 1989;86:3833–3837.
802. Persic L, Roberts A, Wilton J, Cattaneo A, Bradbury A, Hoogenboom HR. An integrated vector system for the eukaryotic expression of antibodies or their fragments after selection from phage display libraries. *Gene* 1997;187:9–18.
803. Caron PC, Laird W, Co MS, Avdalovic NM, Queen C, Scheinberg DA. Engineered humanized dimeric forms of IgG are more effective antibodies. *J Exp* Med 1992;176:1191–1195.
804. Pollack SJ, Jacobs JW, Schultz PG. Selective chemical catalysis by an antibody. *Science* 1986;234:1570–1573.
805. Tramontano A, Janda KD, Lerner RA. Catalytic antibodies. *Science* 1986;234: 1566–1570.
806. Gibbs RA, Posner BA, Filpula DR, et al. Construction and characterization of a single-chain catalytic antibody. *Proc Natl Acad Sci USA* 1991;88:4001–4004.
807. Jackson DY, Prudent JR, Baldwin EP, Schultz PG. A mutagenesis study of a catalytic antibody. *Proc Natl Acad Sci USA* 1991;88:58–62.
808. Tang Y, Hicks JB, Hilvert D. In vivo catalysis of a metabolically essential reaction by an antibody. *Proc Natl Acad Sci USA* 1991;88:8784–8786.
809. Schnee JM, Runge MS, Matsueda GR, et al. Construction and expression of a recombinant antibody-targeted plasminogen activator. *Proc Natl Acad Sci USA* 1987;84:6904–6908.
810. Thrush GR, Lark LR, Clinchy BC, Vitetta ES. Immunotoxins: An update. *Annu Rev Immunol* 1996;14:49–71.
811. Staerz UD, Bevan MJ. Hybrid hybridoma producing a bispecific monoclonal antibody that can focus effector T-cell activity. *Proc Natl Acad Sci USA* 1986;83: 1453–1457.
812. Suresh MR, Cuello AC, Milstein C. Advantages of bispecific hybridomas in one-step immunocytochemistry and immunoassays. *Proc Natl Acad Sci USA* 1986; 83:7989–7993.
813. Mallender WD, Voss EJ. Construction, expression, and activity of a bivalent bispecific single-chain antibody. *J Biol Chem* 1994;269:199–206.
814. Staunton DE, Ockenhouse CF, Springer TA. Soluble intercellular adhesion molecule 1-immunoglobulin G1 immunoadhesin mediates phagocytosis of malaria-infected erythrocytes. *J Exp Med* 1992;176:1471–1476.
815. Martin S, Casasnovas JM, Staunton DE, Springer TA. Efficient neutralization and disruption of rhinovirus by chimeric ICAM- 1/immunoglobulin molecules. *J Virol* 1993;67:3561–3568.
816. Hasemann CA, Capra JD. High-level production of a functional immunoglobulin heterodimer in a baculovirus expression system. *Proc Natl Acad Sci USA* 1990;87:3942–3946.
817. Ma JK, Hiatt A, Hein M, et al. Generation and assembly of secretory antibodies in plants. *Science* 1995;268:716–719.
818. Mhashilkar AM, Bagley J, Chen SY, Szilvay AM, Helland DG, Marasco WA. Inhibition of HIV-1 Tat-mediated LTR transactivation and HIV-1 infection by anti-Tat single chain intrabodies. *EMBO J* 1995;14:1542–1551.
819. Knight KL, Spieker PH, Kazdin DS, Oi VT. Transgenic rabbits with lymphocytic leukemia induced by the c-myc oncogene fused with the immunoglobulin heavy chain enhancer. *Proc Natl Acad Sci USA* 1988;85:3130–3134.
820. Schmidt EV, Pattengale PK, Weir L, Leder P. Transgenic mice bearing the human c-myc gene activated by an immunoglobulin enhancer: A pre-B-cell lymphoma model. *Proc Natl Acad Sci USA* 1988;85:6047–6051.
821. Maxwell IH, Glode LM, Maxwell F. Expression of the diphtheria toxin A-chain coding sequence under the control of promoters and enhancers from immunoglobulin genes as a means of directing toxicity to B-lymphoid cells. *Cancer Res* 1991;51:4299–4304.
822. Pascual V, Capra JD. Human immunoglobulin heavy-chain variable region genes: Organization, polymorphism, and expression. *Adv Immunol* 1991;49:1–74.
823. Miwa H, Nosaka T, Kita K, et al. Immunogenotypes of lymphoid malignancies: The rearrangement of T cell receptor beta chain gene can occur before the gamma chain gene rearrangement. *Jpn J Cancer Res* 1988;79:484–490.
824. Max EE, Maizel JJ, Leder P. The nucleotide sequence of a 5.5-kilobase DNA segment containing the mouse kappa immunoglobulin J and C region genes. *J Biol Chem* 1981;256:5116–5120.